A SIMULATION OF SELECTED OFFENSIVE STRATEGIES IN COLLEGE FOOTBALL

A THESIS

Presented to

The Faculty of the Graduate Division

by

Bruce David Fitzgerald

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Industrial Engineering

Georgia Institute of Technology

April, 1968

A SIMULATION OF SELECTED OFFENSIVE

STRATEGIES IN COLLEGE FOOTBALL

Approved:

____

Chairman

____

____

Date approved by Chairman: April 29, 1968

ACKNOWLEDGMENTS

I wish to thank my advisor, Professor Cecil G. Johnson, for the understanding which he showed during the course of my thesis research and writing. His abilities to guide an undisciplined researcher and to clarify the subtle distinctions between the essential and the ceremonial served to motivate, instruct, and to reveal the excitement of scientific research.

Dr. Jamie J. Goode provided valuable assistance in formulating and statistical testing of hypotheses. His ability to penetrate quickly to the core of a problem provided an understanding of what were reasonable goals for this thesis and how they could be attained.

I wish also to acknowledge the help of Dr. William W. Hines in exposing sloppy writing and fuzzy thinking, and of the Georgia Tech Athletic Association in allowing the use of their files as a source for the data which were necessary for this research.

FOREWORD

Is simulation applicable to college football and would it be a helpful tool for the strategic management of a college football team? Answering these questions is the purpose of this investigation. But before beginning the development of the thesis, a point which is implicit in its contents should be explicitly stated.

The purpose of the thesis is to simulate a system, not to gather data. To build the simulation model, certain data will be gathered, grouped, and tested, and from the results of the tests, inferences made about the game of football. The more data which are gathered, the more (and more powerful) are the inferences which can be made.

College football is the object of extensive record keeping so that the amount of data which a researcher could gather is limited only by the amount of time which he can invest to satisfy the purposes of his research. If his purpose is to find accurate estimates of all the statistical parameters of college football, the researcher must collect many more data than if he needs only close approximations of those parameters.

This thesis is concerned with the estimation of stochastic parameters in college football only because they are some of the elements which must be included in the simulation model. Consequently, low variance in the estimation of the parameters is not necessary. It would be an added luxury but its price, in terms of the amount of data which would have to be gathered, is greater than its potential value to this investigation.

The probability estimates which will be made in the study and used in the simulation model can easily be changed and the simulation rerun if further research into college football gives more precise results.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS. ii

FOREWORD iv

LIST OF TABLES vi

LIST OF ILLUSTRATIONS. vii

LIST OF TESTS. x

SUMMARY. xi

Chapter

I. INTRODUCTION. 1

Empirical and Scientific Decision Methods
Simulation
Literature Survey
The College Football "Industry"
Statement of the Specific Problem
Scope and Limitations

II. Construction of the Model 16

Data Collection
Fixed Parameters
Variable Parameters
Random Number Generation
Summary

III. TESTING THE MODEL 49

Decision Method
The Simulation Experiment
Statistical Tests

IV. CONCLUSIONS AND RECOMMENDATIONS 58

Conclusions
Recommendations

APPENDIX 63

BIBLIOGRAPHY 117

LIST OF TABLES

Table		Page
1.	Opponents and Scores of Georgia Tech in the 12 Games Selected for Study	18
2.	Field Goal Attempts and Results	30
3.	Estimates of Play Timings Used in the Simulation Model	36
4.	Distance from Last Line of Scrimmage that Ball Is Put into Play by the Intercepting Team	38
5.	Yardage and Possession Lost or Gained on Fumbles	39
6.	Penalties when Georgia Tech is on Offense	41
7.	Penalties when Georgia Tech is on Defense	42
8.	Summary of Penalty Data from Tables 6 and 7 Displaying Data According to Whether the Penalty is Against the Offense or the Defense	43
9.	A Display of Data Gathered from the Simulation Experiment and Used in Nonparametric Tests	54
10.	A Display of the Statistics Used and the Results of the Parametric Tests	57
11.	A Display of the Numerical Data Used in the Regression Analysis of Georgia Tech and Opponents' Punting	66

LIST OF ILLUSTRATIONS

Figure | Page

1. A Display of the Data and Regression Equation for Georgia Tech's Punting on First, Second and Third Downs 21
2. A Display of the Data and Regression Equation for Georgia Tech's Punting on Fourth Down 22
3. A Display of the Data and Regression Equation for Opponents' Punting on First, Second and Third Downs . . . 23
4. A Display of the Data and Regression Equation for Opponents' Punting on Fourth Down 24
5. Regression Lines Used in the Simulation Model to Assess the Results of Punting. 27
6. A Display of Kickoff Data 29
7. Probabilities of Successfully Kicking a Field Goal from Selected Points on the Field. 32
8. Field Goal Probabilities as a Function of the Distance of the Scrimmage Line from the Goal Line. 33
9. A Graphical Display of the Ranks of the Actual Game Scores Within Their Simulation Populations Plotted Within 80%, 90% and 95% Critical Regions. 55
10. Listing of the COBOL Simulation Model 79
11. Input to the Simulation Model for the Georgia Tech-Florida Game, 1960 92
12. Input to the Simulation Model for the Georgia Tech-Alabama Game, 1960 93
13. Input to the Simulation Model for the Georgia Tech-Tulane Game, 1960. 94
14. Input to the Simulation Model for the Georgia Tech-Southern California Game, 1961 95
15. Input to the Simulation Model for the Georgia Tech-Tulane Game, 1961. 96

Figure | Page

16. Input to the Simulation Model for the Georgia Tech-Rice Game, 1961. 97

17. Input to the Simulation Model for the Georgia Tech-Tulane Game, 1962. 98

18. Input to the Simulation Model for the Georgia Tech-Auburn Game, 1962. 99

19. Input to the Simulation Model for the Georgia Tech-Florida Game, 1963 100

20. Input to the Simulation Model for the Georgia Tech-Florida State Game, 1963 101

21. Input to the Simulation Model for the Georgia Tech-Navy Game, 1964. 102

22. Input to the Simulation Model for the Georgia Tech-Georgia Game, 1964 103

23. Output of the Simulation of the Georgia Tech-Florida Game, 1960 104

24. Output of the Simulation of the Georgia Tech-Alabama Game, 1960 105

25. Output of the Simulation of the Georgia Tech-Tulane Game, 1960. 106

26. Output of the Simulation of the Georgia Tech-Southern California Game, 1961 107

27. Output of the Simulation of the Georgia Tech-Tulane Game, 1961. 108

28. Output of the Simulation of the Georgia Tech-Rice Game, 1961. 109

29. Output of the Simulation of the Georgia Tech-Tulane Game, 1962. 110

30. Output of the Simulation of the Georgia Tech-Auburn Game, 1962 111

31. Output of the Simulation of the Georgia Tech-Florida Game, 1963 112

Figure Page

32. Output of the Simulation of the Georgia Tech-Florida State Game, 1963 113

33. Output of the Simulation of the Georgia Tech-Navy Game, 1964. 114

34. Output of the Simulation of the Georgia Tech-Georgia Game, 1964 115

35. Output of the Georgia Tech-Florida Game of 1960 with Georgia Tech's Decision Rule Modified so that Georgia Tech Does Not Kick on Third Down. 116

LIST OF TESTS

Test | Page

1. Test of the Hypothesis that the Variances about the Four Punting Regression Lines are Equal 68

2. Test of the Hypothesis that the Variances about the Two Regression Lines for Georgia Tech Punting are Equal 69

3. Test of the Hypothesis that the Slopes of the Two Punting Regression Lines for Georgia Tech are Equal. . 70

4. Test of the Hypothesis that the Two Regression Lines for Georgia Tech Punting are Concurrent 71

5. Test of the Hypothesis that the Variances about the Two Regression Lines for Opponents' Punting are Equal. 72

6. Test of the Hypothesis that the Slopes of the Two Regression Lines for Opponents' Punting are Equal 73

7. Test of the Hypothesis that the Two Regression Lines for Opponents' Punting are Concurrent 74

8. Test of the Hypothesis that the Variances in the Populations of Kickoffs by Georgia Tech and by Opponents are Equal. 75

9. Test of the Hypothesis that the Means of Populations of Kickoffs by Georgia Tech and by Opponents are Equal 76

10. Test of the Hypothesis that the Probability of a Pass Interception by Georgia Tech is Equal to the Probability of a Pass Interception by Opponents. 77

11. Test of the Hypothesis that the Variance about 463 Population Means is Equal to the Variance about 66 Population Means. 78

SUMMARY

This thesis is an investigation of the application of simulation to the problem of assessment of offensive strategies in college football. An offensive strategy is a set of decision rules which specifies which of four strategy alternatives (pass, run, punt, field goal) will be chosen under any possible state of the relevant decision variables. At present, alternative strategies are assessed chiefly through the use of empirical methods or in actual games by the coach making decisions in accordance with his strategy and then running the plays and observing the outcome.

In this investigation, data are gathered from 12 Georgia Tech games played during the years 1960-1964. From these data necessary parameters are found for a simulation model. The parameters are of two types. Fixed parameters are those which remain constant in the simulation model regardless of the teams which are being simulated. New variable parameters are introduced for each team to be simulated.

The simulation model is solely a device to assess strategy. It does not deal with how a strategy should be chosen but with providing a quantitative assessment of a given strategy. To test the simulation model a simple decision procedure which approximates the actual strategies used by each of the coaches in the games in the data sample is designed. Using this decision procedure and input parameters which are estimated statistically from the game data, each of the 12 games is simulated 25 times.

The populations of simulated scores are compared with the actual game scores and the hypothesis that the actual score is randomly drawn from the population of simulated scores is tested using both nonparametric and parametric tests.

For the nonparametric test the hypothesis is rejected for three of the 12 simulated games at the 80 per cent level of confidence, for one at the 90 per cent level and for one at the 95 per cent level. These results may be ascribed to the level of the test.

The condition for the parametric test, bivariate normality, is not met by the data so that while the results of the tests are not as favorable as the nonparametric results, they do not give sufficient cause to overrule the conclusion implied by the nonparametric tests: that the simulation model does produce a population of scores from which the real game score might have been randomly selected. This is the objective of the simulation model.

The fact that a scientific study of football strategy selection is needed is emphasized by the disparity between a coach's beliefs regarding the effectiveness of third down punting and the actual data which show third down punts to be no more effective than fourth down punts.

It is concluded that simulation can be an effective method of assessment of football strategies but that further research should expand the model of this thesis to consider defensive strategies and more strategy options before it will be of use to the coaching community.

CHAPTER I

INTRODUCTION

The proper selection of game strategy by the management of a college football team is essential to the consistent attainment of the team's objective: winning football games. While a strong team may overwhelm a decidedly weaker opponent regardless of the strategy used by either team, the coaches of two teams of equal skills, abilities, strengths and weaknesses must seek an advantage in strategy selection to break the deadlock of tactical equality. Two coaches in such a position may be likened to opposing chess players. At the game's outset each has equal physical resources: The game will be won by the superior strategist, he who can better manage his own resources and can take advantage of his opponent's errors.

Football coaches have generally been quick to apply new technologies to the purpose of winning football games but the pertinent literature carries no accounts of their having used any scientific procedures to aid them in the important area of strategy selection. This is a report of an investigation of the feasibility of using computer simulation as a device for assessing alternate offensive strategies and it is a foundation for further research into the application of scientific decision-making techniques to college football.

Before developing the problem in more specific terms, some necessary background information is presented. This includes brief discussion of scientific (opposed to empirical) decision methods, simulation as a tool for analyzing decisions, and a survey of previous literature which is relevant to this study's objectives.

Empirical and Scientific Decision Methods

The industrial engineer bases his decisions upon knowledge of relationships between variables which is acquired by the scientific method of observation, hypothesis formulation, and testing. The process of decision-making does not demand that scientific methods be used. Empirical methods, especially when the decision-maker has extensive experience with similar decisions can yield the "right" decision. It should not be accepted without proof that a given decision can better be handled by scientific than by empirical methods.

A paradigm for scientific decision models and its representation in the problem of offensive strategy selection are:

(1) a decision-maker who has the problem and must make the final choice (in football, the head coach and his staff),

(2) a set of options which are the possible courses of action (the alternative offensive strategies under consideration),

(3) a problem context or "states of nature" which consists of those factors outside the control of the decision-maker which affect the outcome of a given course of action (in football, the action of the opponent and certain randomly occurring events such as fumbles, penalties, and incomplete passes),

(4) an outcome-payoff relationship which assigns to each course of action under all specific states of nature some relevant outcome to

which the payoff can be computed (the yardage gained or lost, points scored, and change of possession on each play), and

(5) an element of doubt or uncertainty regarding the "best" course of action (which gives rise to the second-guessings of the Monday-morning quarterbacks and, more importantly, a need for a means of reducing this element of uncertainty).

A large variety of scientific methods have been applied toward this end by industrial engineers. These include mathematical programming, game theory, queueing theory, statistical inference, decision theory, and simulation.

Simulation

Simulation is an important method of analyzing a system which is so complex as to be impossible or impractical to describe in terms of a set of mathematical equations. It is by no means a recent innovation nor is it the sole property of the scientific community.

> . . . The oldest and most familiar examples of simulation are ordinary physical models used in the crafts. The seamstress has her dress form, the apprentice barber his dummy head. The infantry soldier, unable to stick his bayonet into real people until the proper time, practices his art on simulated human torsos. (14, p. 125.)

Digital computer simulation has proven to be an effective means of generating numerical data describing processes which otherwise would reveal such information at a high cost if at all. It has been used with both deterministic and stochastic processes.

Stochastic simulation allows for the element of chance in addition to asserting specific relationships among the variables. Using available data, probability distributions of a system's relevant random variables can be estimated. In the course of the simulation the values of the variables are assigned according to the probability distributions by random number generation.

A technique closely related to simulation is gaming. Often in the literature "simulation" and "gaming" are discussed together and are sometimes used interchangeably. In this paper they will be used in accordance with Rapoport's distinctions:

> The term "simulation", in my opinion, ought to be reserved to refer to procedures in which both the assessment of the situation and the decisions are carried out in accordance with formal rules. Since the decision rules are often complex and, also, assessments require a great deal of calculation, the computer takes over both these functions; but they could also, in principle, be performed by human beings instructed to follow the rules of assessment and decision to the letter.
>
> "Gaming," on the other hand, is an appropriate term for simulated situations in which either assessment or decisions but not both are made more or less freely by human beings. (14, p. 129.)

The designers of simulation and gaming models agree that the principal value of these procedures is in furthering the development of theory and as teaching aids. Many business management games and war games, based upon computer assessment of the results of decisions, have been designed to give neophyte decision-makers experience with the complexities of the policies which they will later have to analyze, choose among, and implement. This study could provide similar games for those concerned with decision-making in college football.

As an aid to the furthering of the development of theory, simulation offers conditional results of the sort: If the decisions were guided by strategy x_i and the assessments of the situations from the decisions were determined in some manner, f(X), then the results y_i, where $y_i = f(x_i)$, will follow. In this study, X is the set of all possible strategies, f represents the mathematical simulation model which this study will develop and which assigns to each strategy, x_i, one and only one y_i, a sample from a population of possible outcomes. The strength of the results of the simulation model and the generalizations which can be inferred from the results will necessarily depend upon how accurately the simulation model, f(X), duplicates the real process of a football game and upon how closely X represents the set of strategies available to the decision-maker.

Essentially, simulation is a method of representing a physical system by a mathematical system. By manipulating the mathematical equivalent, insight can be gained into how the physical system would react to equivalent manipulations. The system can be either deterministic or stochastic. An appropriate model for college football is stochastic.

By representing the actual game by mathematical relationships (building a model) and varying the relationships to represent alternate strategies, the strategies can be assessed on paper. If the relationships are properly conceived to approximate correctly the game of football, then the outcome of the real game may be inferred from the outcome of the simulation model.

Literature Survey

Tocher (17, p. 2-6) assigns the origins of mathematical simulation to three groups of scientific workers: mathematical statisticians, applied mathematicians, and operations analysts.

In the early nineteenth century mathematical statisticians, concerned with the difficulties of describing a population given only a sample from it, designed sampling experiments to give experimental verification of and confidence in their methods. A close approximation to a probability distribution was created, samples were taken, and the resulting frequency chart of sampled values was compared with the predictions of theory.

Applied mathematicians became interested in simulation as a method of solving problems involving partial differential equations. Von Neumann and Ulam conceived during the Second World War of formulating and solving some of the equations as random walks. One of the simplest and most powerful applications of this idea was the evaluation of a multidimensional integral. They called the technique "Monte Carlo."

Operations research developed during the time of the Second World War by building models of the systems it studied and using these models to gain insight and quantitative information about the systems. The models studied were primarily probabilistic and the conditions and restrictions could rarely be incorporated in models that mathematical methods could solve so the scientists turned to a technique of experimentation: simulation.

After the Second World War operations analysts began applying their methods to the problems of industry and government operations outside the military establishment. These applications have generated an extensive literature. *Bibliography on Simulation* (19) cites 948 papers, articles, and books dealing with simulation during the period 1960 through 1964.

Simulation methods are popular with economists because of the difficulty of running experiments in economic systems. The nature of the systems rarely allows economists the freedom of other scientists to change some variables intentionally to see what happens to others. Simulation allows this freedom.

Two significant macrodynamic models are those of Duesenberry, et al. (5), and Orcutt, et al. (13). Duesenberry's model contains over 250 equations and represents a composite of the theories and empirical findings of a large number of American economists. Orcutt's model was constructed in terms of microcomponents, individuals and families, rather than aggregative components.

One of the best known and most widely used simulation techniques is Jay Forrester's *Industrial Dynamics* which he defines as

> . . . the study of the information feedback characteristics of industrial activity to show how organizational structure, amplification (in policies), and time delays (in decisions and actions) interact to influence the success of the enterprise. It treats the interactions between the flows of information, money, orders, materials, personnel, and capital equipment in a company, an industry, or a national economy. (6, p. 13.)

Other areas in which simulation has been applied include traffic control, communications, distribution systems, education, military logistics, manufacturing, psychology, queueing, and management games.

The first of the modern management games was developed in 1956 by the American Management Association and is documented in *Top Management Decision Simulation* by Ricciardi, et al. (15). The game was used as part of a training course held by AMA and it relied upon an IBM 650 computer to perform the necessary computations.

In 1958 the first major non-computer-based management game was reported in an article in *Harvard Business Review* (1). The game became widely known as a result of this article and has had extensive use due to its low cost and its translation into many languages.

Since these early developments in management gaming, games have been developed and used by such organizations as Rand Corporation, Carnegie Institute of Technology, IBM, Procter and Gamble, and the Department of Defense. The extensive directory in *Management Games* (10) lists 113 games in use at the time of the book's publication. It includes games in such diverse areas as insurance underwriting, materials management, bank management, tire sales, oil exploration, supermarket management, international relations, research and development, and airline operations.

There are no reports in the literature of applications of simulation techniques to college football. Two significant studies have been made of professional baseball, though neither used simulation techniques in its analysis.

Earnshaw Cook performed a detailed statistical analysis of professional baseball in *Percentage Baseball* (4). After studying over 750,000 times at bat over 10 seasons of major league play and using elementary probability theory in his analysis, Cook concluded that sacrifice bunts are ordinarily useless, batting lineups are sloppily ordered, the platooning of personnel is grossly miscalculated, pitchers should never bat, and intentional walks should not be given. Jim Brosnan, a professional baseball pitcher for 17 years, said in a review of Cook's book,

> The average manager directs a baseball game according to the BOOK, an unwritten compendium of personal experiences, abstracted folk lore, and generalized statistics. . . . Cook calls the BOOK a collection of dusty myths and whimsical managerial inspiration. (2, p. 112.)

Ira Horowitz presented in the *Journal of Industrial Engineering* (9) a decision-making model for coaches to use in baseball planning and player trading decisions. Horowitz explained that professional baseball is of interest to the industrial engineer because

> . . . whether it is looked on as sport, business, or a combination of the two, organized baseball is run by managements that are attempting to maximize something, be it profit, utility, or number of first place finishes. . . . this model provides a framework for decision-making that can be extended beyond this particular "industry." (9, p. 170.)

In addition to these two studies of baseball, the use of scientific methods of football scouting is reported by W. N. Wallace in (18). Designed by Computer Applications Inc. of Silver Spring, Maryland, the system was originally used by the University of Maryland to simplify the process of analyzing an opponent's game films.

> In preparing to play their next foe, all serious football teams at the college and professional level attempt to take from films the probabilities--called frequencies--of the opponent's offense and defense. Football teams are creatures of habit and they like to do what they know they can do best.
>
> The computer has taken much pain out of the extraction of frequencies. A film need only be run through once with the proper information extracted and programmed. Into a computer it goes and four hours later a book dissecting a team's offense--and defense if requested--is available with the frequencies coded on wide sheets. (18)

Since the introduction of the system at the University of Maryland three years ago, similar systems have been adopted by seven professional football teams. They have also adopted another system which processes the mass of information on 1500 college players scouted by the professionals each year.

The College Football "Industry"

For the same reasons professional baseball is of interest to Horowitz (9), college football is a potential area of study for the industrial engineer, and the large amounts of data which Cook found necessary for the success of his study of baseball are available in the form of play-by-play descriptions of each game which can be supplemented by films if the researcher needs more detailed data.

College football has become a large and profitable business in the United States. A team rated highly in either the Associated Press or the United Press International polls of sports writers and football coaches may anticipate near-capacity crowds every week. The difference between a filled and a half-filled stadium can be more than $150,000 to the athletic associations of the two schools involved in the game.

Both for financial considerations and the personal and social satisfaction of fielding a winning football team, it behooves a team's coaches and managers to investigate every possible method of improving their won-lost record. Athletic associations spend large sums for player scholarships, recruitment, training, conditioning, housing, and boarding toward this end.

As new technologies are developed they have been applied by football coaches to their special management objective: amassing a greater number of points than their opponents. Motion pictures are made of each game. Video tape units are used to give players instant and clearer understanding of their performances. Telephones and small radio sets are used to afford the strategy decision-makers better communication. However, none of the new technologies in the field of management science has been significantly applied to strategy selection in college football. This is a report of an investigation of the assessment of selected offensive strategies in college football through the use of simulation.

Statement of the Specific Problem

On each play the offensive team must make a decision among four alternatives: run, pass, punt, and attempt a field goal. This is the offensive strategy decision and it is the ultimate responsibility of the team's coach. The coach must consider the characteristics of each type of play: a punt surrenders possession of the ball and the offensive thrust to the opponent; a successful field goal represents an immediate three points while the unsuccessful effort may be less effective than a punt; a pass has greater probability of a long yardage gain than does a run but it also has a greater risk of losing possession of the ball. In addition there are certain randomly occurring events to be considered: blocked punts, penalties, fumbles. Other factors which must be accounted for in the decision process are down, game score, time remaining in the game, field position, yardage needed for a first down, and the results of the preceding plays.

After the strategy decision is made a specific play must be chosen. The considerations are no longer of a general strategic nature. The specific plays are considered on the expected performances of individual players, both offensive and defensive. At this decision level the coach is no longer a strategist but a tactician. Occasionally there are overlappings between the strategic and the tactical decision levels but treating the strategic decision independently of the tactical is justified by the infrequency of these overlaps.

At present the strategy decision is made by empirical methods. The coach uses experience and conjecture to guide his decision and test his analysis. The decision must be made quickly but with consideration for all the factors mentioned above. Computer simulation may provide an effective decision-making aid for the coach.

The week before a game the coach studies films and scouting reports of the opposition. From them he can formulate opinions of strengths and weaknesses and finally his own strategy, a set of rules that specifies which strategy decision will be made under any possible state of the relevant decision variables. But in choosing between alternate strategies the coach has only empirical methods of assessing the effectiveness of each. Computer simulation can provide quantitative measures of the effectiveness of alternative strategies.

From repeated simulations of each strategy, populations of scores are obtained. Each score represents one completed game, one possible outcome of the game if the coaches act according to the given strategies and if the teams perform according to the assumed input parameters.

These populations of simulated scores can yield quantitative comparisons among several alternative offensive strategies. Assuming each simulated outcome is equally likely to occur,

$$\text{probability of winning} = \frac{\text{number of favorable outcomes}}{\text{total number of simulations}} .$$

After simulation populations and probability estimates are obtained for each decision strategy under consideration, the strategy which gives the greatest likelihood of achieving management's primary objective--scoring more points than the opposition--can be selected by the coach.

The strategies which can be simulated may be as complicated as the decision-maker wishes to define them. The simplest class of strategies dictates the same decision, a pass or a run, on every play. A more complicated class of strategies is one which calls for a different decision on each down, such as, "pass on first down, run on second, pass on third, punt on fourth." More complicated classes of strategies are generated by adding more conditional clauses: "Punt on fourth down unless you are within 40 yards of the opponent's goal line. If you need less than five yards for a first down and are more than three points behind, run. Otherwise, try a field goal." Any strategy which accurately describes the empirical decision process which the coach uses would necessarily have a large number of qualifying clauses. The simulation model would allow him to try strategies which he could never experiment with in a real game such as the first class of strategies mentioned above.

The two distinct problems of any simulation model are assessment and decision-making. In an actual game, assessment of a decision is accomplished by running the chosen play. In a simulation model it is accomplished through the use of mathematical methods. The objective of this investigation is to find the necessary mathematical relationships to assess the results of a given offensive strategy.

The first half of this thesis is a report of the analysis of data gathered from 12 football games played during the 1960-64 period. The objective of this analysis is to define the input parameters necessary to represent each team's performance and to find the frequency of occurrence of random events such as fumbles, penalties, and pass interceptions.

After the simulation model is constructed, it is tested by simulating the games which were analyzed in the first half of the study. A crude decision-making method which approximates the decisions made by the coaches during each of the games is used in simulating each game several times. The populations of simulated scores are compared with the actual scores as a measure of the effectiveness of the simulation model.

Scope and Limitations

The simulation model is intended neither to predict a game's actual score nor to prescribe how a team should play a game. The model will assess the probable outcomes of a given strategy but will not propose a better strategy. For any proposed strategy, it will reveal a population of possible scores, and, if the model is satisfactory, it is reasonable to assume that the actual game score will be a member of this population. Too many random occurrences--penalties, pass interceptions, fumbles--are involved in the final outcome for a reliable prediction of the score, or for the simulated scores even to be expected to have a small variance.

Because of the dearth of previously reported research for this investigation to build upon, this will serve as an exploratory study into the application of science to decision-making in college football. Its object is not to provide an immediate aid to the football-coaching community but to see if aid can be given by the techniques of management science. Simulation seems to be the most readily applicable technique and is therefore the method used in this investigation.

In constructing the simulation model, some assumptions will be made which are justified by the nature of this initial research but which should be tested before any final solutions are proposed. In some statistical populations normal distributions are assumed because of the statistical tests which are allowed under this assumption. A more rigorous study would include tests of many of these assumptions. In other cases, statements concerning estimates of parameters could be strengthened by more data, and some populations which must be combined for this study could be treated separately.

An assumption which should be relaxed in more detailed studies is that the outcome of the play is controlled by the offensive strategy decision but not by the defense's expectations concerning the offense's plans. Intuitively it seems that a defensive team which correctly predicted, and set its defenses for, each type of play that the offense ran should be expected to stop the offense shorter than they would if they had no idea of what type of play would be run.

CHAPTER II

CONSTRUCTION OF THE MODEL

Assessment of a given offensive strategy in an actual football game is accomplished by making the decisions in accordance with the strategy, running the plays, and watching the results. On each play fumbles either occur or they do not, passes are either completed or not, yardage is either gained or lost and is always in some measurable amount. In a simulation, each possible outcome of the real-world game must be a possible outcome of the simulated game, and the simulation model must have some way of injecting such random occurrences as fumbles, pass interceptions, and the amount of yardage gained or lost on each play.

This will be accomplished by estimating the probability of each outcome of the real-world game, and then assigning the outcome according to the probability estimates by the process of random number generation. Gathering data, using them to find the appropriate probability estimates to include in the simulation model, and random number generation are the topics of this chapter.

Two types of parameters will be included in the simulation model. Fixed parameters will remain constant for every game which is simulated and variable parameters will be changed at the beginning of every game. Whether a given parameter is to be fixed or variable in the model will be determined by the amount of data which are available to estimate the parameter.

Data Collection

For the purposes of this study 12 games were randomly selected from those played by Georgia Tech during the five-year period 1960-1964. The games selected are listed in Table 1.

Georgia Tech was chosen as a data source because of the convenience of obtaining data from a local college and because of the willingness of the Athletic Association to allow the use of its files. The years 1960-1964 were chosen as a period of time in which the rules of the game (20) remained relatively constant but of sufficient length that different players would prevent the data from being biased by the characteristics of a few special players or teams. Twelve games were chosen so that data could be obtained from 2000 separate plays. Slightly fewer than that number were actually in the games studied.

The Georgia Tech Athletic Association has on file written play-by-play descriptions and slow-motion films of each game which was selected for this study. For all the data which are needed, the play-by-play summaries are adequate. They list the yard line on which each play began, the yard line on which it ended, the players handling the ball, the type of play, and the player making the tackle.

The only ambiguity in this source of data was that when a quarterback was listed as the ball carrier it is not always clear whether he was running the ball or was actually trying to pass and was unable to find any receivers in the open. The ambiguity was resolved arbitrarily by classifying all gains by the quarterback as runs and all

Table 1. Opponents and Scores of Georgia Tech in the 12 Games Selected for Study

Year	Opponent	Tech's Score	Opponent's Score
1960	Florida	17	18
	Alabama	15	16
	Tulane	14	6
1961	Southern California	27	7
	Tulane	35	0
	Rice	24	0
1962	Tulane	42	12
	Auburn	14	17
1963	Florida	9	0
	Florida State	15	7
1964	Navy	17	0
	Georgia	0	7

losses as pass negatives (plays on which the offensive strategy decision was to pass but the defense tackled the passer for a loss before he could throw the ball).

The data which were gathered on each play are: game, quarter, team with ball, down, type of play, yardage gained or lost on the play, and other information regarding the occurrence or nonoccurrence of fumbles, pass interceptions, and penalties. These data were entered into punched cards for analysis utilizing the Burroughs B-5500 at the Rich Electronic Computer Center. These data provide all the information which will be needed for this study. They are not shown in this thesis but are on file with the School of Industrial Engineering.

Fixed Parameters

The fixed parameters are composed of two types. The first are those parameters which must be considered fixed because an inadequate sample size makes it impossible to subgroup them by team. This type includes kicking parameters (punts, quick kicks, kickoffs, field goals), and points after touchdown (deciding whether to go for one or two points, probability of being successful). In this study it is a necessary, though questionable, assumption that these parameters remain constant from team to team, but to eliminate the assumption would require a much larger data sample than is feasible for this investigation. Consequently, data on these parameters are subdivided no further than the two groups: Georgia Tech and opponents. In studying some of the parameters they are not subdivided at all.

The second class of fixed parameters are those which cannot be varied regardless of the number of observations. These are play timings, pass interceptions, fumbles (loss of yardage, likelihood of recovery), and penalties (distance, team penalized). These parameters may be regarded as characteristics of the game of football.

Punting

To facilitate the handling of punting data, two events which could be treated independently have been combined: the distance which the ball is kicked in the air from the line of scrimmage, and the distance which the ball is run back by the receiver. The combining of these events is justified by the fact that the football coach, in deciding whether or not to punt, is more concerned with where the ball will finally come to rest at the end of the play than he is with what intermediate points it may visit during the play.

Punting data were divided into four groups and a regression line was fitted to each. The groups are: Georgia Tech, first, second, and third down punts; Georgia Tech, fourth down punts; opponents, first, second, and third down punts; and opponents, fourth down punts. Punting data and the regression lines are displayed in Figures 1-4. Numerical data which were extracted from these figures are shown in the Appendix, Table 11.

These groupings were made in order to test whether punts on third down are actually more effective, as many football coaches and writers maintain, than fourth down punts, and to test whether Georgia Tech is a more effective punting team than its opponents.

For both Georgia Tech and the opponents, the fourth down data were further partitioned into two groups, punts kicked between the zero and 40-yard lines and punts which were kicked from outside the 40-yard line. This was done to afford a better comparison between third and fourth down data since third down punts never occurred outside a kicking

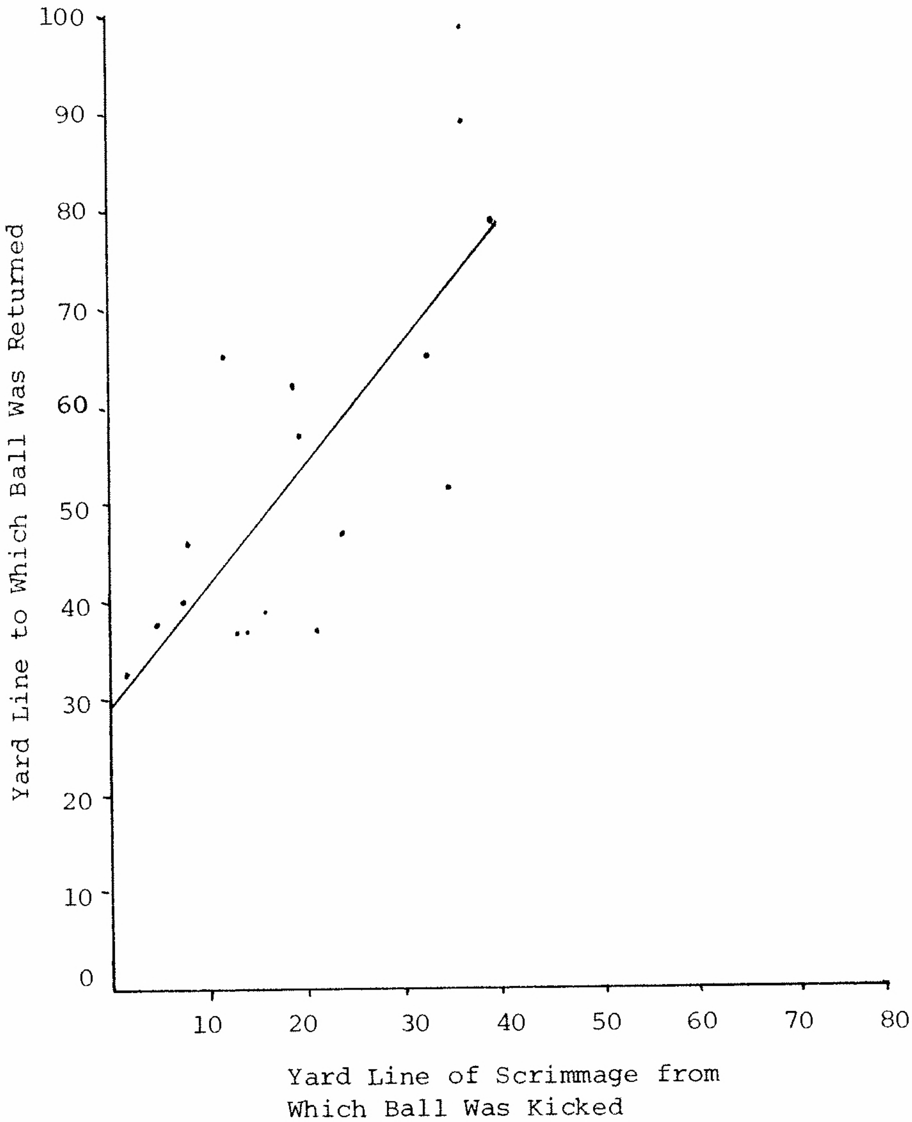


Figure 1. A Display of the Data and Regression Equation for Georgia Tech's Punting on First, Second and Third Downs

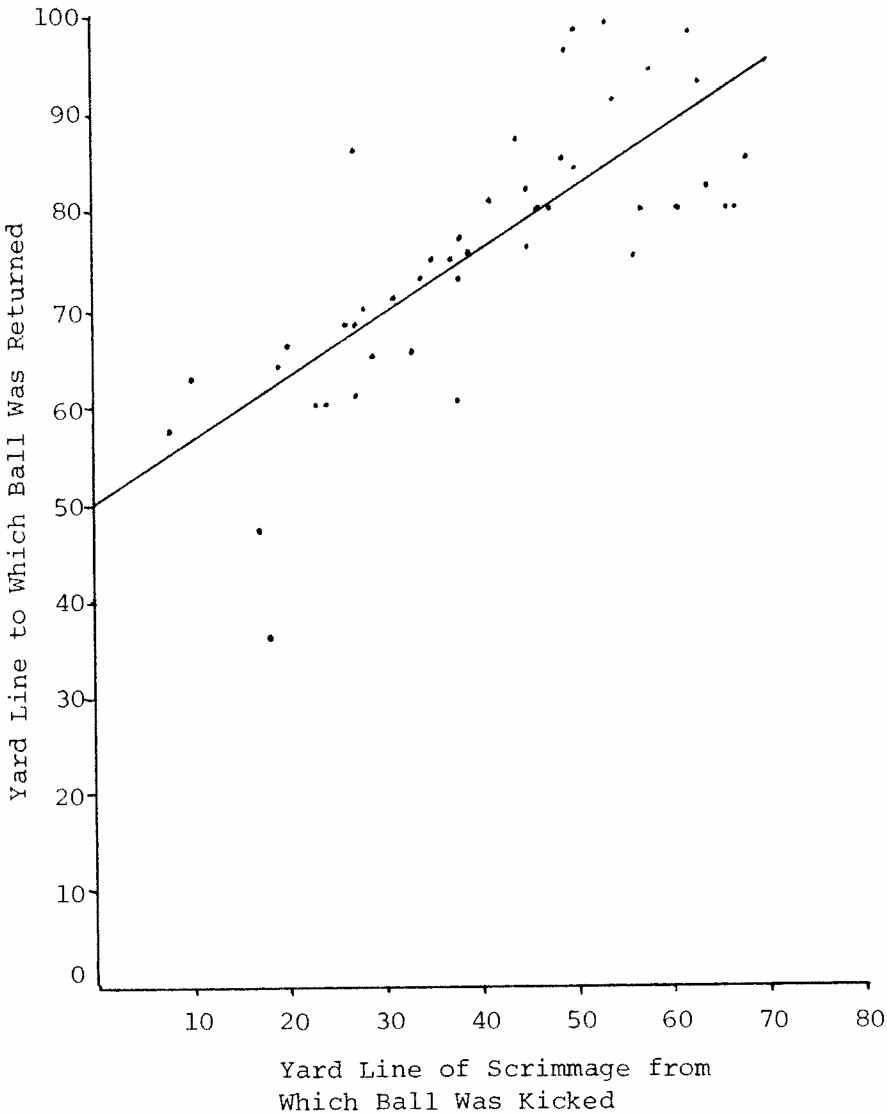


Figure 2. A Display of the Data and Regression Equation for Georgia Tech's Punting on Fourth Down

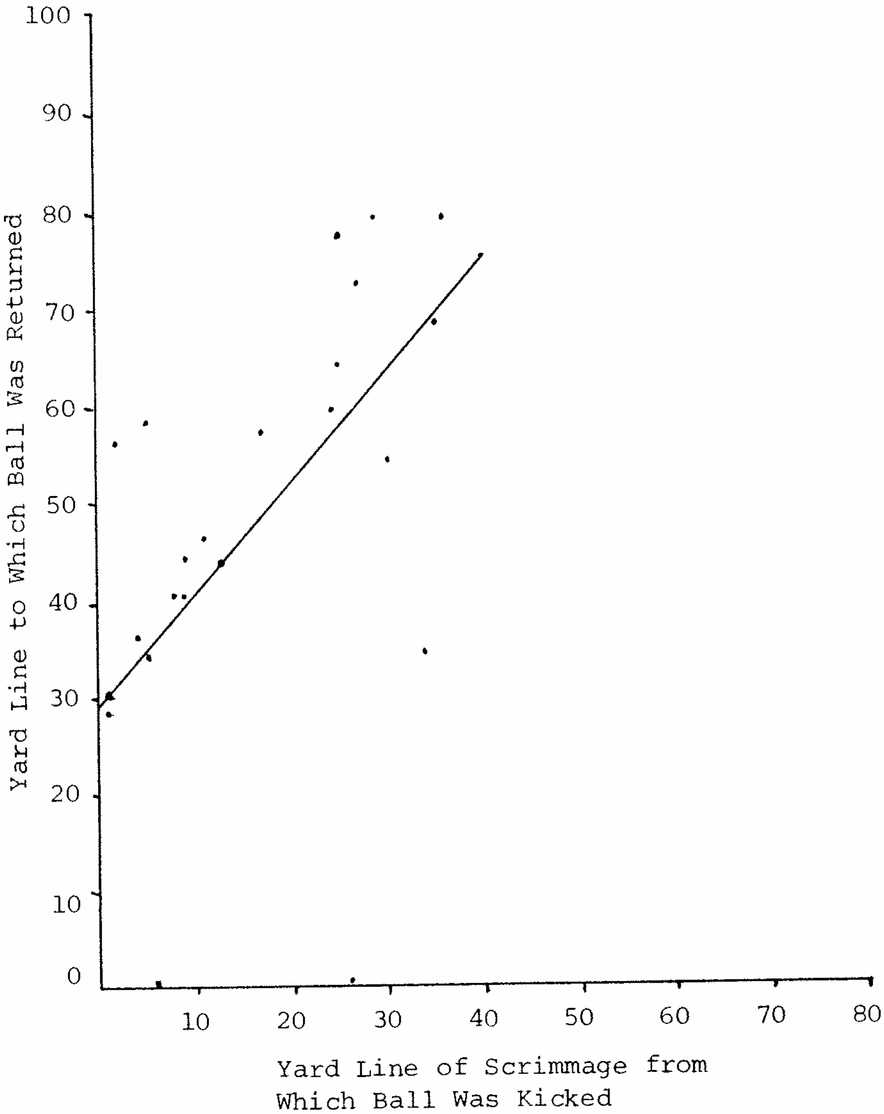


Figure 3. A Display of the Data and Regression Equation for Opponents' Punting on First, Second, and Third Downs

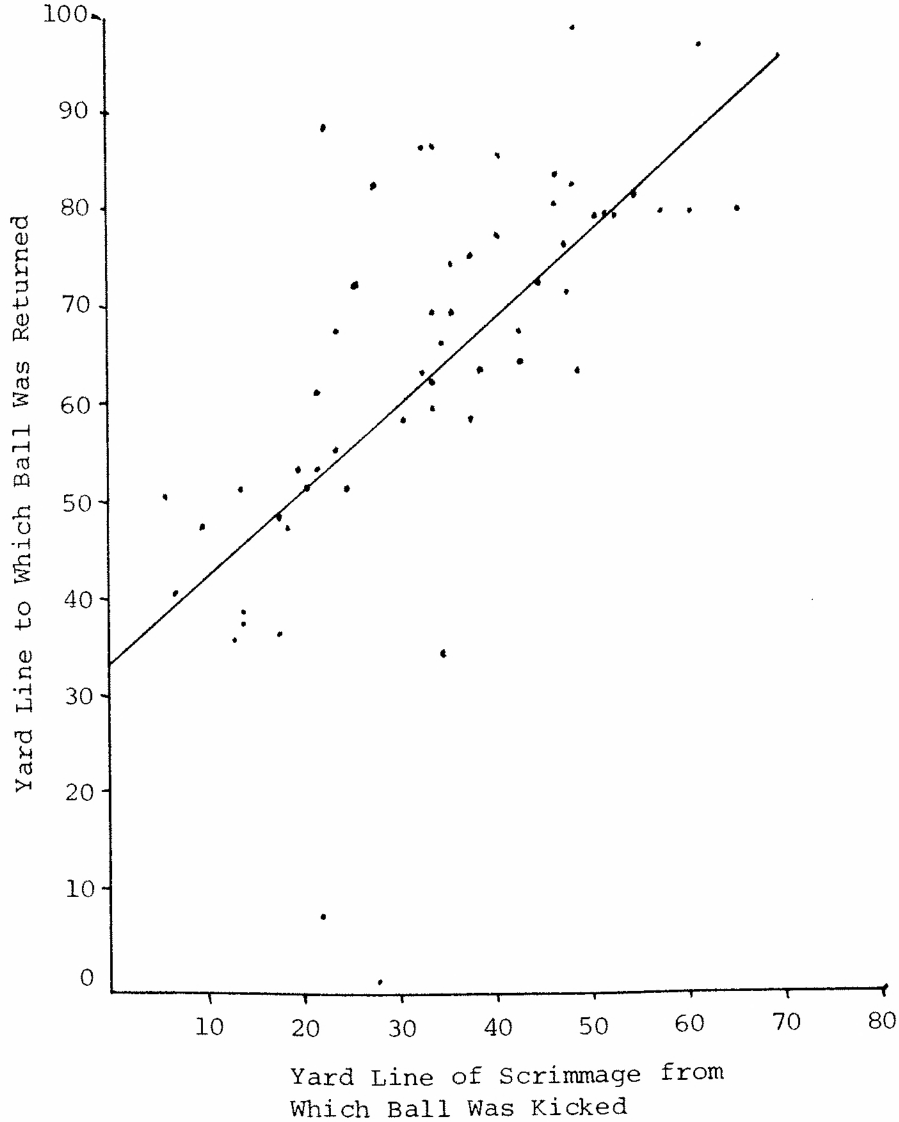


Figure 4. A Display of the Data and Regression Equation for Opponents' Punting on Fourth Down

team's 40-yard line.

To test the hypothesis that the four regression lines can be represented by a single line, the procedure which is used demands that: the hypothesis of equality of variances about the lines must not be rejected, equality of slopes must not be rejected, and finally the hypothesis that variance about the four lines is equal to the variance about a single line must not be rejected.

The hypothesis that the variances about the four lines are equal was rejected (Test 1 in the Appendix) so that no single line could be used to replace all four lines. The variances about the two lines for the opponents were found to be significantly larger than the variances about the two lines for Georgia Tech. This would indicate that the punting skills of 12 opposing teams varied more from team to team than did Georgia Tech's from game to game.

The next set of hypotheses tested were that the two regression lines for Georgia Tech can be combined into a single line (Tests 2-4) and the two regression lines for the opponents can be combined into one line (Tests 5-7).

No significant differences were found between the variances or the slopes in either case and finally the hypothesis that the four regression lines could be represented by two, one for Georgia Tech punting inside its 40-yard line and one for the opponents punting inside their 40-yard line, was not rejected.

To estimate the effect of a punt outside the 40-yard line, an additional constraint was placed on those regression lines. In order to avoid a discontinuity at the 40-yard line, the regression line for punts outside the 40-yard line was required to intersect the regression line for punts inside the 40-yard line at the 40-yard line.

The regression lines which will be used in the simulation model to estimate the most likely result of a punt are illustrated in Figure 5. The variances about these regression lines are listed in Table 11 in the Appendix.

Blocked Punts, Penalties and Fumbles on Punts

On the 141 punts in the data sample, there were four occasions on which the punt was fumbled by the receiver and recovered by the kicking team; two punts which were blocked by the receiving team, one of which was recovered by the kicker, one by the blocker; five penalties, three against the defensive team and two against the offense.

These data will be used in the simulation model in the following ways:

(1) There is a probability of .03 that the punt will be fumbled by the receiver and recovered by the kicking team.

(2) There is a probability of .02 of a punt's being blocked and an equal likelihood of its being recovered by either team ten yards behind the line of scrimmage.

(3) There is a probability of .02 of a 15-yard penalty against the defense.

(4) There is a probability of .01 of a 5-yard penalty against the defense.

(5) There is a probability of .02 of a 5-yard penalty against the offense.

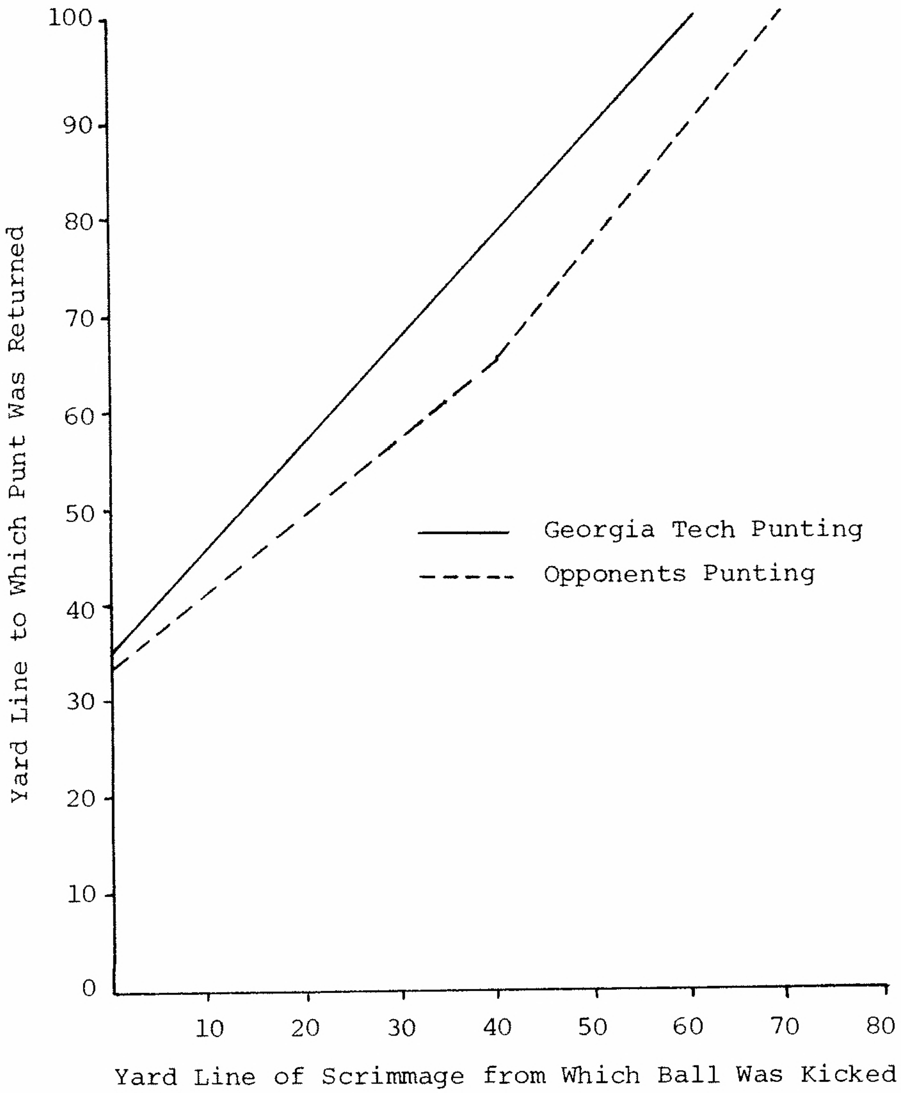


Figure 5. Regression Lines Used in the Simulation Model to Assess the Results of Punting

Quick Kicks

Quick kicks, punts which are executed in such a manner as not to indicate to the opposing team that a punt is planned, were observed only seven times in the games which were studied. Because of this lack of data, they are not separately analyzed in this study.

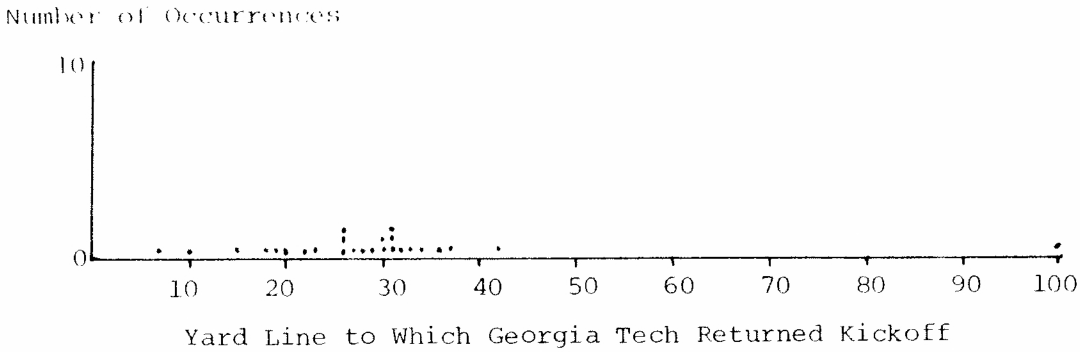
Quick kicks seem to be an element of college football which is periodically in vogue. The frequency of occurrence of quick kicks seems to depend upon the rules in effect concerning the kicking team. In the simulation model, no quick kicks will be allowed and all punts will come from the populations considered in the previous section.

Kickoffs

As in the punting data, two separate events are combined in the kickoff data: the distance the ball is kicked in the air and the distance that the ball is run back by the receiver. Of interest to this model is where the ball was put into play after the kickoff rather than whether it got there due to the strength of the kicking or the receiving team. Thus combining these two events is a suitable way of handling them.

The data are displayed in Figure 6. They are divided into two groups, kickoffs by Georgia Tech, and kickoffs by opponents.

The two samples were tested for equality of variances and means. A highly significant difference was found between the variances, and, using the Behrens-Fisher test, no significant difference was found between the means. The tests are shown as Tests 7 and 8 in the Appendix.



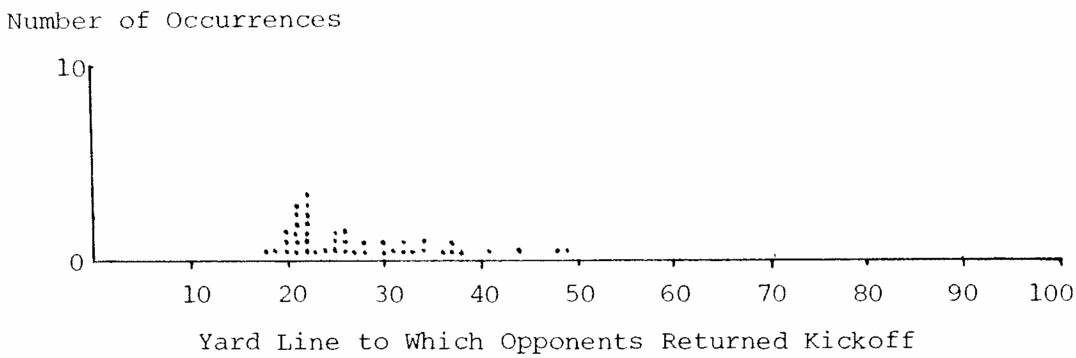


Figure 6. A Display of Kickoff Data

As a result of the tests, in the simulation model the yard line on which the ball is placed after kickoffs to Georgia Tech will come from the normal population with mean 28.7 yards and standard deviation 17.6 yards. For the opponents, the mean is 28.7 yards and the standard deviation, 7.9 yards.

Field Goal Kicking

In the 12-game data sample, Georgia Tech was successful in seven of nine field goal attempts, the opponents were successful in five of nine. The results of the field goal attempts are summarized in Table 2.

Table 2. Field Goal Attempts and Results

	Yard Line Kicked From	Result
Georgia Tech	09	Good
	11	Good
	18	Good
	18	Good
	18	Good
	23	Good
	27	Bad
	30	Good
	35	Bad
Opponents	05	Good
	06	Good
	09	Good
	11	Bad
	12	Bad
	18	Good
	25	Good
	28	Bad
	35	Bad

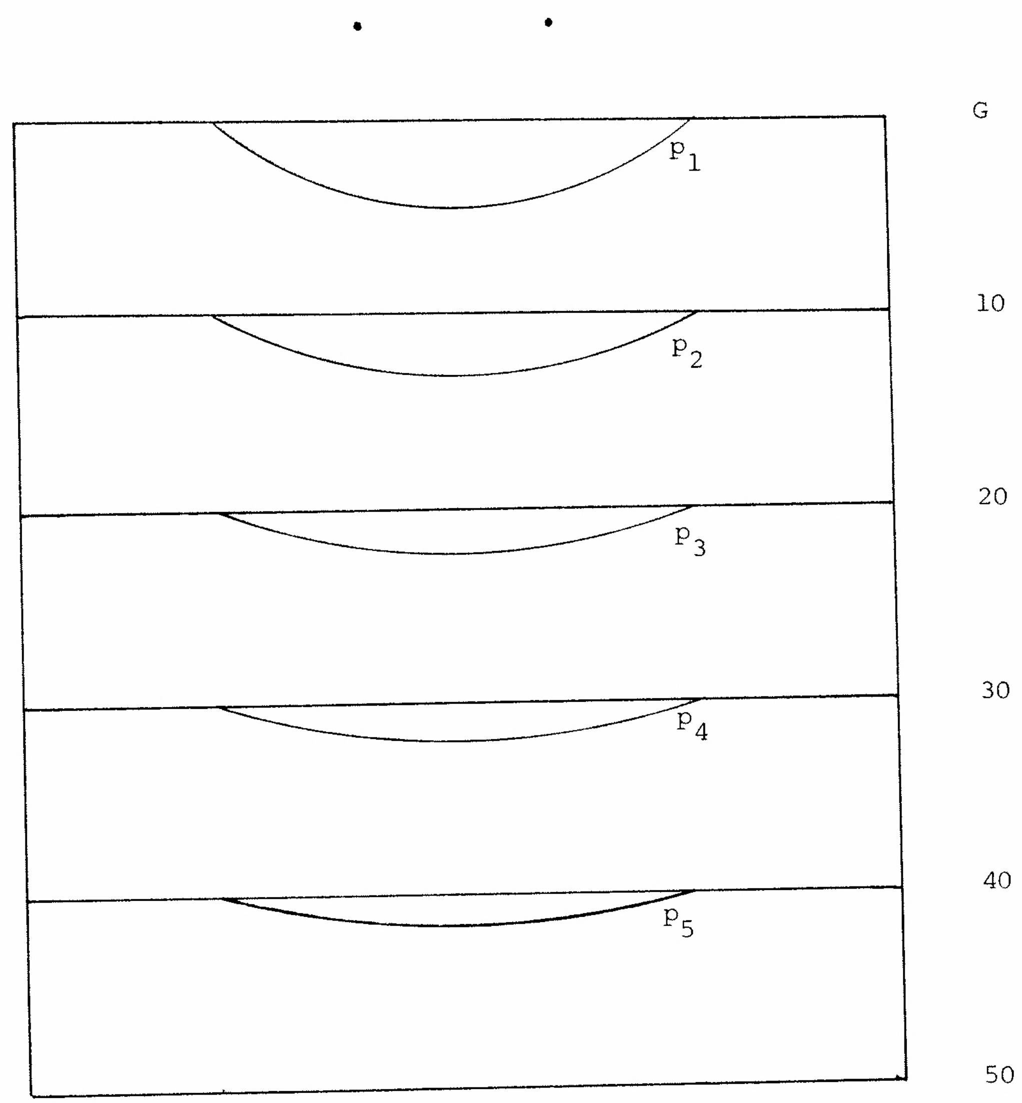
This amount of data is too small to make any strong statements about the probability of kicking a field goal from any point on the field. Further data could be gathered, but to fit them to the simulation model of this paper would oversimplify the problems of field goal kicking.

The difficulty of kicking a field goal is a function of the distance from the goal line, and the angle from which the kicker faces the crossbars. A kick from the five-yard line at an extreme angle may be more difficult than one from the 15-yard line with the ball positioned in the center of the field. A correct and complete probability model for field goal kicking should have approximately the form illustrated in Figure 7.

Any attempt to describe the surface in Figure 7 would be useless to this paper since the simulation model will not consider the effects of the distance of the ball from the sidelines upon offensive strategy.

Since the true surface in Figure 7 would be difficult to derive and of limited value when found, it is reasonable for the purposes of this study to use a cruder approximation of the probabilities of kicking field goals. Figure 8 shows the approximation which is used. No attempt will be made to create separate curves for Georgia Tech and the opponents.

A straight line passing through (3, .85) and (35, .05) (see Figure 8) was chosen. The first point was selected because the attempts for points after touchdowns, which are essentially field goals from the three-yard line, were successful in 29 of 34 attempts by all teams in the 12 games. The second point was chosen with consideration



Note: $p_1 > p_2 > \cdots > p_5$

Figure 7. Probabilities of Successfully Kicking a Field Goal from Selected Points on the Field

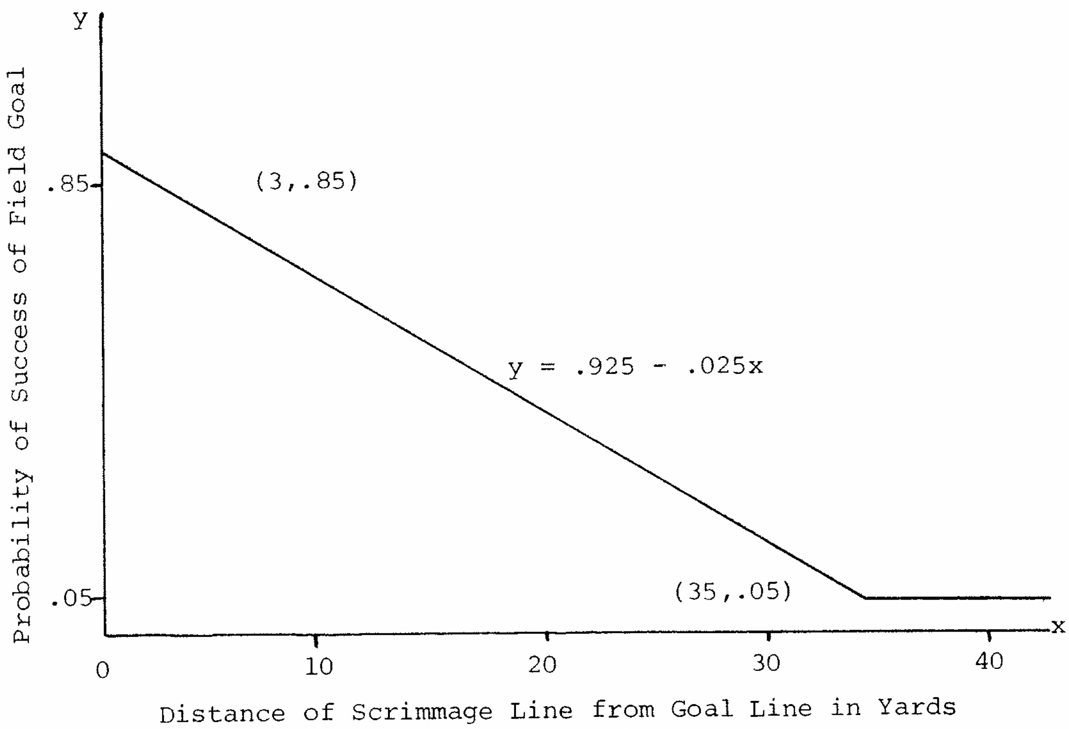


Figure 8. Field Goal Probabilities as a Function of the Distance of the Scrimmage Line from the Goal Line

of the fact that no field goals were tried from beyond the 35-yard line. This indicates the small probability of success which coaches assign to kicks from that distance.

Equally good arguments could be advanced for other curves; however, this curve is adequate for the purposes of this study. In practice a coach has many more observations of his kicker in actual game and practice conditions upon which to base his decision. He could and should construct a more accurate probability curve.

Points after Touchdowns

After a touchdown is scored in college football the scoring team is given the option of trying for one or two extra points. One point is scored by kicking a field goal from the three-yard line, two points by either a completed pass or a run over the goal line from the three-yard line.

Formulating a decision rule for trying for one or two points is in itself a difficult problem having as its basis the utilities of winning, tying and losing a game. To avoid the problem of trying to describe the logic which a coach uses in deciding whether to go for one or two points and of estimating the probabilities of success from a small amount of data, the following analysis will detail the method used in the simulation model to assess the number of points which each touchdown is worth.

The 12-game data sample showed (combining Georgia Tech and opponents) 2 eight-point touchdowns, 28 seven-point touchdowns and 13 six-point touchdowns. This gives probability estimates of .05 for eight-point touchdowns, .65 for seven-point touchdowns, and .30 for six-point touchdowns. After each touchdown a two-digit random number will be generated. If it is less than 05 the touchdown will count eight points, if it is less than 70 the touchdown will count seven points, otherwise it will count six points.

Times of Plays

An important factor in offensive strategy selection is the amount of time remaining in the game. As the losing team tries to conserve the final minutes of the game, a trade-off develops between time and field position. An incomplete pass may be more valuable to the losing coach than a pass completed for only short yardage because the incomplete pass stops the clock and time is not wasted while the offensive team is in the huddle.

Any simulation model of football must have some method of associating an amount of elapsed time with each play which is run. Any "acceptable" method must meet the criteria of: preventing the simulation model from allowing an inordinately large number of plays in a simulated game; and reflecting the relative differences in the amount of time consumed by the different types of plays.

Two methods which satisfy the above criteria were considered, multiple regression and stopwatch time study. Multiple regression was chosen for use in this study.

In each quarter there are run a variable number of each type of play, but the total amount of time consumed in each quarter is always 900 seconds. For each of the four quarters in each of the 12 games studied, it is known how many of each type of play was run. Using multiple regression it is possible to associate with each of the eight play types (run, pass, incomplete pass, pass negative, punt, field goal, quick kick, and kickoff) a constant coefficient which represents the amount of time that can be attributed to the play type. This coefficient will include both the time used in running the play and the time consumed between plays. It allows for time used by times out and saves the simulation model from having to determine when to call times out.

A standard multiple regression program from the Burroughs B-5500 library was used to obtain the estimates of play times shown in Table 3.

Table 3. Estimates of Play Timings Used in the Simulation Model

Type of Play	Time of Play (in Seconds)
Run	32
Pass	26
Pass Negative	28
Incomplete Pass	-3
Punt	30
Field Goal	5
Quick Kick	90
Kickoff	15

These estimates include two peculiar results: the times for quick kicks and incomplete passes. The large estimate of time for quick kicks is caused by the small number of occurrences of this play and is not crucial to the simulation model since quick kick is not one of the strategy alternatives allowed to the offensive teams.

The negative time estimate for incomplete passes is reasonable since, in stopping the clock, an incomplete pass saves the time which a team normally takes to huddle for the next play. Since the total amount of time that it takes to huddle for and run an incomplete pass is three seconds less than the time saved by the clock's not running during the following huddle, it is reasonable to say that the net effect is to give the team an extra three seconds on the following play.

The coefficients obtained from multiple regression meet both criteria of acceptability and will be used in the simulation model.

A stopwatch time study was considered. The films of past Georgia Tech games could be viewed to find the time consumed by each of the eight types of plays considered. However, other elements of the game which are absorbed by multiple regression, the amount of time needed to huddle, times out, and the effects of penalties, could not be analyzed from the films. Since these are important elements of play timing, and they are included in the time estimates obtained by multiple regression, a stopwatch time study was not conducted.

Pass Interceptions

Data on pass interceptions are displayed in Table 4. Georgia Tech intercepted 8 of the opponents' 254 passes. The opponents intercepted 14 of Georgia Tech's 241 passes. From these data, the probability that a Georgia Tech pass will be intercepted is estimated as $\theta_1 = .058$, the probability that a pass decision by the opponents will result in an interception is $\theta_2 = .031$.

Table 4. Distance from Last Line of Scrimmage that Ball Is Put into Play by the Intercepting Team

Interceptions by Georgia Tech	Interceptions by Opponents
-23 (Touchdown)	17
2	25
14	2
-6	33
-6	38
31	4
13	0
13	36
	2
	5
	8
	31
	9
	15

NOTE: Negative yardage means it was closer to the passing team's goal line by the amount shown. Positive means it was closer to the intercepting team's goal line.

The null hypothesis that $\theta_1 = \theta_2 = \theta$ was tested (Test 10 in the Appendix) and was not rejected at the 5 per cent level of significance. Thus, in the simulation model the probability that a pass is intercepted will be $\theta = 22/495 = .044$ for both teams.

The effect of a pass interception in the simulation model is assessed as follows: The team intercepting takes possession at a point x yards from the last line of scrimmage where x is randomly chosen from the population with mean 11.9 yards and standard deviation 15.5 yards. These parameters were estimated by combining the data on pass interceptions for Georgia Tech and the opponents and taking maximum likelihood estimates from them.

Fumbles

Data on occurrences and outcomes of fumbles are displayed in Table 5.

Table 5. Yardage and Possession Lost or Gained on Fumbles

Fumbles by Georgia Tech: Lost	Fumbles by Georgia Tech: Recovered	Fumbles by Opponents: Lost	Fumbles by Opponents: Recovered
-1	-3	-3	-3
-2	-5	2	6
5	0	2	-5
-3	6	-2	2
0	7	-1	14
0	0	1	0
1	5	-7	5
-3	1	-4	1
-3	1		-8
-1	-5		-9
-3			22
-4			-5
-9			
-13			
-1			
-2			

Because fumbles often occur before it is obvious whether the play is a pass or a run, it is not possible to make any statements as to whether the probabilities of fumbling on the two types of plays are the same or different. It will be assumed in this study that they are the same.

The small number of occurrences of fumbles in a game gives no clear trends as to whether Georgia Tech fumbles more often than its opponents or whether Georgia Tech was more likely to recover fumbles than its opponents.

The following assumptions were made and incorporated in the simulation model without testing: Fumbles are equally likely to be recovered by the offense and the defense. The probability of a fumble on a pass or a running play is constant from game to game and team to team and can be estimated by dividing the total number of runs and passes in the data sample into the total number of fumbles (46/1648 = .028). The population of yardage gained on fumbles is normally distributed and has a mean of -.6 yards and standard deviation of 5.7 yards, the parameters estimated using data in Table 5.

Penalties

Of the 1602 running and pass plays in the data sample, there was a penalty against one of the two teams on 63 of the plays. This gives an estimate of .039 for the probability of a penalty on a pass or a run play.

The data were grouped only by offense and defense rather than by play type or by Georgia Tech and opponents. This is because the small number of penalties occurring in any game would make statistical testing of hypotheses difficult.

The data on penalties are displayed in Tables 6 and 7.

Table 6. Penalties when Georgia Tech is on Offense

Pass Play		Running Play	
Against*	Distance	Against*	Distance
2	11	1	5
2	5	2	5
2	5	1	13
2	13	1	5
1	5	1	15
1	15	1	5
1	5	2	15
2	10	1	5
2	5	1	5
2	5	1	5
		1	11
		2	5
		2	4
		1	5
		1	5
		2	15
		1	12
		1	5
		2	13
		1	15
		1	5
		1	13

*1 represents Georgia Tech. 2 represents the opponents.

Of the 62 penalties, 39 were against the offensive team giving the probability estimate of .629 that if a penalty occurs it will be against the offense and .371 it is against the defense. These two estimates are used in the simulation model.

Given that a penalty occurs and that it is known which team it is against, the next problem is to assess the amount of yardage of the penalty.

Table 7. Penalties when Georgia Tech is on Defense

Pass Play		Running Play	
Against*	Distance	Against*	Distance
2	5	1	15
2	5	1	15
2	5	2	14
1	15	2	5
2	5	2	5
1	23	1	15
2	16	1	5
1	5	2	5
		2	5
		1	5
		2	29
		2	10
		1	5
		2	5
		2	5
		2	8
		2	5
		2	5

*1 represents Georgia Tech, 2 represents opponents.

Table 8 is a summary of the penalty data of Tables 6 and 7.

Most penalties are either 5 or 15 yards and are usually assessed from the line of scrimmage. The fact that some are assessed from the place of infraction accounts for those penalties in Table 8 which are not for 5 or 15 yards. Also, no team may be penalized more than half the distance to its own goal line.

After an analysis of the data in Table 8, the following methods were chosen to assess the yardage of penalties. If the penalty is against the offensive team, the probability of a five-yard penalty is

Table 8. Summary of Penalty Data from Tables 6 and 7 Displaying Data According to Whether the Penalty is Against the Offense or the Defense

Yards	Number of Occurrences
Penalties Against Offense	
5	25
8	1
10	1
11	1
12	1
13	2
14	1
15	5
16	1
29	1
Penalties Against Defense	
4	1
5	10
10	1
11	1
13	2
15	8
23	1

.625, the probability of a 15-yard penalty is .125, and there is a probability of .250 that the yardage will fall randomly in the interval [0,20]. If the penalty is against the defensive team, the probability of a five-yard penalty is .40, the probability of a 15-yard penalty is .35 and there is a probability of .15 that the yardage will fall randomly in the interval [0,33].

The only problem remaining after the penalty has been assessed is whether it will be accepted or rejected. The logic which a coach uses in accepting or rejecting a penalty is difficult to duplicate.

The approximation of the logic which will be used in this model is detailed below.

If the penalty is against the offensive team, the following conditional statements will be examined until one is found to which the conditions are met.

(1) If the offensive team lost more yardage on the play than they would on the penalty, reject the penalty.

(2) If it was fourth down and insufficient yardage for either a first down or a touchdown was gained, reject the penalty.

(3) If the ball is inside the defense's ten-yard line, accept the penalty.

(4) If the offense gained a first down, accept the penalty.

(5) If the gain plus the amount of the penalty is greater than seven yards, accept the penalty. Otherwise, reject the penalty.

If the penalty is against the defensive team, the following conditional statements will be examined until one is found to which the conditions are met.

(1) If the penalty is greater than the gain on the play, accept the penalty.

(2) If there was a touchdown scored on the play, reject the penalty.

(3) If there was a first down made on the play, reject the penalty.

(4) If the length of the penalty plus three yards is greater then the gain, accept the penalty. Otherwise, reject the penalty.

Variable Parameters

In the preceding section the assessment of the results of two of the offensive strategy decisions, punting and field goal kicking, was discussed, and the random, unexpected outcomes of offensive strategy decisions were analyzed. In this section the method of assessing the amount of yardage lost or gained as a result of the decision to pass or to run will be discussed.

Since passing and running are the offensive strategy decisions which are made most often, to give a more accurate simulation it is desirable that the parameters which are used in assessing their results be changed for each game which is simulated. The passing and running parameters represent the interaction of both teams: the strengths and weaknesses of each team's offense in playing against the other's defense.

If a decision is made to pass (and assuming there are no fumbles or penalties) the coach may anticipate one of four outcomes. The passer may be tackled before the ball is thrown (a pass negative), the pass may be incomplete, the pass may be intercepted, or the pass may be completed. The problem of interceptions has been treated as a fixed parameter. The probabilities of a pass's being either not thrown (pass negative) or incomplete will be variable parameters and will be estimated by dividing the pass data for each team in a game into three groups, passes negative, incomplete passes, and completed passes. By dividing the number of passes negative by the total number of pass plays, an estimate of the probability that a pass is not thrown and that the passer is thrown for a loss, is obtained.

Assuming that the pass is thrown, then it will either be incomplete, intercepted, or completed. The probability that a pass is incomplete is estimated by dividing the number of incomplete passes by the total number of passes thrown.

After the outcome of the decision to pass or to run is known, the problem of assessing the yardage lost or gained remains. Since each of the three outcomes, run, pass negative, and completed pass occurs frequently in every game, the data may be grouped into several populations. A grouping into as many as 96 populations per game is not unreasonable. It would classify yardage gained first by team, then by type (run, pass negative, and pass complete), then by quarter, and finally by down. (This gives 96 populations: 2 (teams) x 3 (types) x 4 (quarters) x 4 (downs) = 96). The parameters of each of these populations may be estimated for each game by statistical methods. However, since there are only about 130 plays in a game, many of which are punts and field goals, the number of observations of each population will necessarily be small. Also, a simulation model with such a large number of input parameters would be useless to a football coach who would have to estimate--before the game is played--each of the parameters.

The problem would be greatly simplified if the 96 populations could in some way be pooled. Since it seems intuitively necessary that teams and play types be kept separate, the smallest number of populations which seems feasible is six: for each team, passes negative, runs, and complete passes.

Since there are 12 games to be simulated, each containing as many as 96 populations, the hypothesis that the 1152 populations could be represented by 72 populations (six in each game) was tested. (Because many of the populations had no observations, the theoretical problem was reduced to testing the hypothesis that 463 populations could be represented by 66 populations.) It was assumed that each population is normally distributed, and, using an F test, the hypothesis was not rejected at the 5 per cent level of significance. The test is shown in the Appendix (Test 11). Consequently the variable input parameters for each team in each game to be simulated are: probability that a pass is not thrown, probability that a thrown pass is incomplete, and the mean and standard deviation of the yardage gained or lost on runs, passes negative, and completed passes.

Random Number Generation

An acceptable method of generating a sequence of random numbers for this simulation study has to meet the criteria that the numbers be uniformly distributed, statistically independent, generated at a high rate of speed and in a minimum amount of computer memory, and non-repeating for the desired sequence length.

The simulation of a complete game of football requires about 620 random three-digit numbers. The final experiment of this investigation, 25 simulations of each of the 12 games listed in Figure 1, makes a sequence of approximately 186,000 desirable.

Discussion of random number generating methods in *Computer Simulation Techniques* (12, pp. 47-57) shows that by suitable choice of a constant, an initial value, and modulus, all the criteria listed above will be satisfied by a multiplicative congruential generator.

The generator selected is:

$$n_{i+1} = a \cdot n_i \pmod{m}$$

where a = 31621, n_o = 173964213, and m = 10^9. This multiplicative procedure will produce 5 x 10^{9-2} = 50,000,000 (12, p. 54) random numbers before repeating.

Summary

This chapter contains the details of the analysis used in studying the elements which must be contained in a simulation model to assess the effects of a given offensive strategy. As a result of the analysis it is possible to combine fixed and variable parameters from 12 past Georgia Tech games and anticipate that the results of the simulations of these games will in some way be similar to the past games.

The actual simulation model is shown as Figure 10 in the Appendix. It is in the form of a computer program written in COBOL for the Burrough's B-5500. The program is a combination of all the elements of college football which were discussed in this chapter arranged in accordance with the game's rules and to permit all the outcomes of the real-world game.

CHAPTER III

TESTING THE MODEL

In this chapter the results of a test of the simulation model which was designed in Chapter II will be reported. Since the simulation model itself is only a device to assess the results of offensive strategy decisions, a scheme for decision-making will be devised to approximate roughly the actual decision-making procedures which were used by the coaches involved in the 12 games in Table 1. These decision parameters and the variable parameters which were found to be necessary in Chapter II will be the inputs to the simulation model which will simulate each of the games 25 times. The resulting populations of scores will be compared with the actual game scores using both parametric and nonparametric statistical methods. This will provide a test of how well the simulation model can assess offensive strategy decisions.

Decision Method

In making his offensive strategy selection, the coach considers the down, game score, time remaining in the game, field position, yardage needed for a first down, and results of the preceding games. A method of making decisions which is not a function of all these variables is subject to criticism as an oversimplification of the coach's decision problem. To put all these variables together in a decision model which approximates the logic which a football coach uses is a formidable task because of the large number of variables affecting his decision, and it is not within the scope of this thesis. It will be avoided in this study by adopting one of the "oversimplified" decisions methods for testing the simulation model.

Each game in the data sample will be examined and the total number of each type of decision which each coach made on each of the four downs will be counted. This counting gives 24 4x4 matrices, the rows of which are the four downs, the columns are the four play types, and the elements, n_{ij}, $(i, j \leq 4)$, are the number of decisions on the ith down which were of the jth play type.

Each of these 24 matrices will be transformed to a probability matrix by dividing each n_{ij} of each matrix by $\sum_{j=1}^{4} n_{ij}$. The 24 new matrices obtained after this transformation will be 4x4 matrices with rows representing downs, columns representing play types and each element, $p_{ij} = \frac{n_{ij}}{\sum_{j=1}^{4} n_{ij}}$, representing the probability of running a play of the jth type on the ith down.

For each game to be simulated there are now two probability matrices, each one representing a crude approximation of the way in which one of the participating teams actually made its offensive strategy decisions. A probability matrix will be used in the following way: On the ith down, a team will have to choose one of the four alternative play types, $j = 1,2,3,4$. A random number, r, will be generated in the interval $[0,1)$. If $r \varepsilon [0,p_1)$, choose play type 1. If $r \varepsilon [p_1, p_1+p_2)$, choose play type 2. If $r \varepsilon [p_1 + p_2, p_1 + p_2 + p_3)$, choose play type 3. If $r\varepsilon[p_1 + p_2 + p_3, 1]$ choose play type 4.

This decision method represents the way the coaches in the real-world game actually made their decisions according to the single variable, down. More sophisticated decision methods could be used, but the purpose of this research is not to study the decision process but to develop a method of assessing a given set of decision rules. The decision rules outlined above are a sufficient set.

The Simulation Experiment

Each of the 12 games was simulated 25 times. Initially each game was simulated only 10 times, but this gave unsatisfactory results in that the output was too sensitive to the sequence of random numbers. With 25 simulations, the output does not vary significantly if the random number sequence is changed.

The inputs to the simulation model, the variable parameters and the decision probability matrices, are shown in the Appendix as Figures 11-22.

It is possible to print play-by-play descriptions of each of the 300 games which were simulated, but they are in themselves of very little value. The only important output of this simulation experiment is the populations of simulated final scores. The final scores were summarized after each game had been simulated 25 times. These summaries are displayed in the Appendix (Figures 23-34, in both graphical and tabular forms.

Statistical Tests

Statistical tests of the results of the simulations are many-one comparisons. The objective of the simulation model is to produce a population of scores from which the real-world score might have come. The null hypothesis in testing the 12 simulations is that the real-world score is a randomly drawn member of the simulation population.

In testing this hypothesis, both parametric and nonparametric statistical methods will be used. The assumption which will be used to make the parametric tests is that the scores have a bivariate normal distribution. This is intuitively a bad assumption because a bivariate normal distribution implies a continuous cumulative density function but football scores are discretely distributed. It also implies a unimodal distribution and a distribution of football scores would be multimodal.

Though the conditions for the bivariate normal tests are not met, the tests will be made and their results reported along with the results of the nonparametric tests. Because the tests under the assumption of bivariate normal distribution are more powerful than the nonparametric tests, the information which they provide can be of some use in assessing the effectiveness of the simulation model when it is observed in conjunction with the nonparametric tests.

Nonparametric Test

The nonparametric test which is used is described in the Appendix, page 65. It is based upon a ranking of each of the 24 sets of simulated scores (two sets in each of the 12 games) and including the actual game's score in the ranking. Under the null hypothesis the actual score is randomly drawn from the simulated scores. Consequently it would be expected that the two ranks would be fairly close to the average rank of the 26 members of the population, 13.5. Expressing the ranks as a two-dimensional point, (r(Georgia Tech), r(opponents)), the probability is found of a point being as far or farther from (13.5, 13.5) than the rankings of the actual game score.

Two assumptions are made which lead to the critical regions shown in Figure 9. The first assumption, that the ranks of the Georgia Tech scores and the ranks of the opponents' scores are independently distributed was tested using the Spearman's Rank Correlation Test (16, pp. 202-213). The region for the coefficient, Spearman's Rho, is $r_s > .465$. From the Spearman's Rho coefficients shown in Table 9, two of the simulated populations (games 8 and 12) were found to be correlated in some way.

The second assumption which was made is that the points representing the ranks are uniformly distributed. The ranks are actually discretely distributed but the process of breaking ties makes ranks such as 1.5 possible. The assumption of uniformity was made to simplify the calculation of critical regions. The resulting critical regions are concentric circles with center at (13.5, 13.5). Critical regions for 80 per cent, 90 per cent, and 95 per cent confidence limits are shown in Figure 9.

The 12 points corresponding to the ranks shown in Table 9 are plotted in Figure 9. Three of the points fall outside the 80 per

Table 9. A Display of Data Gathered from the Simulation Experiment and Used in Nonparametric Tests

Game	Spearman's Rho	Rank of Georgia Tech Score	Rank of Opponent's Score
1	.240	11.50	20.00
2	.450	15.00	15.00
3	.260	24.00	5.50
4	.396	16.50	7.00
5	.086	26.00	1.50
6	.289	19.00	1.00
7	.173	18.50	20.50
8	.677	14.50	11.50
9	.454	10.00	5.50
10	.407	8.00	7.00
11	.170	6.00	5.00
12	.558	2.50	13.50

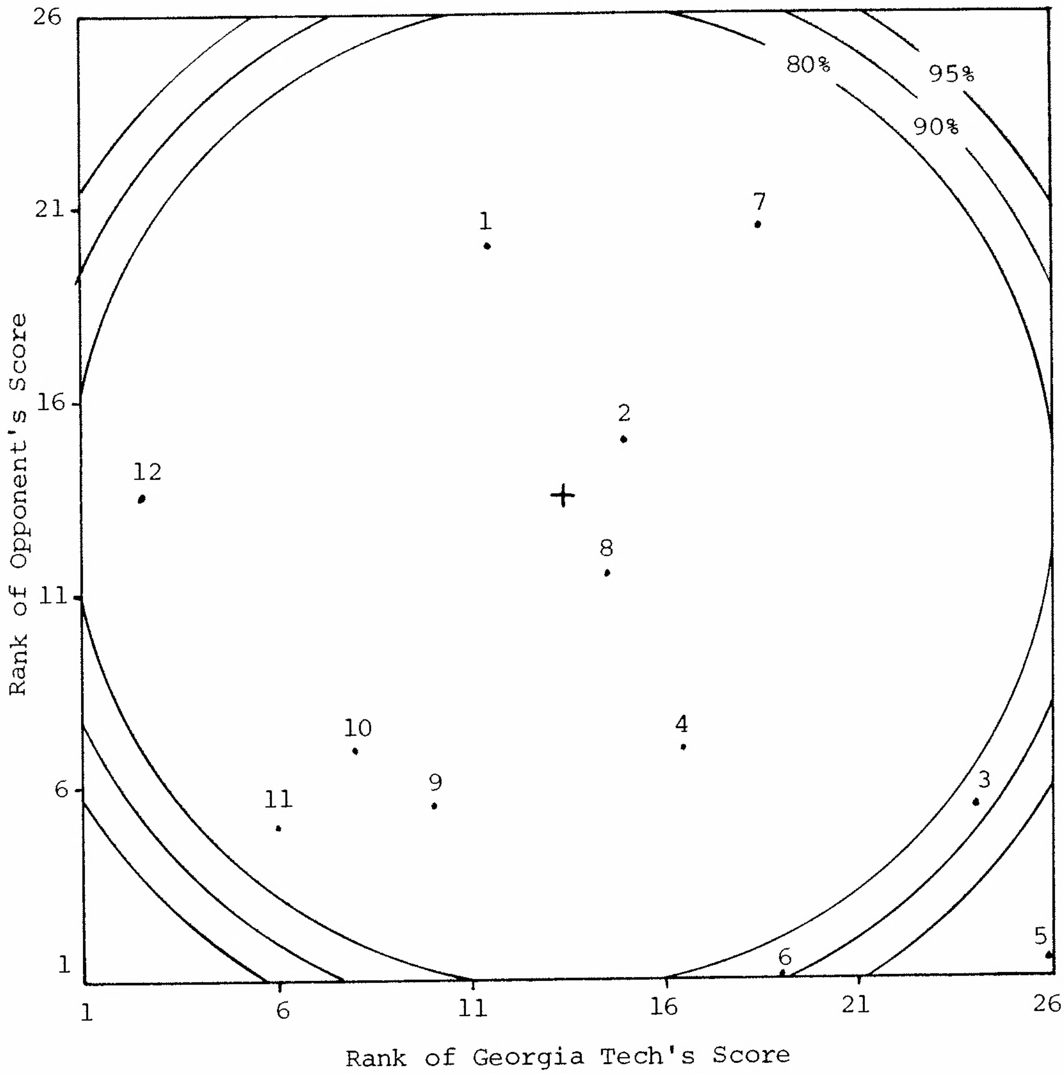


Figure 9. A Graphical Display of the Ranks of the Actual Game Scores Within Their Simulation Populations Plotted Within 80%, 90%, and 95% Critical Regions

cent confidence region, two fall outside the 90 per cent confidence region, and one outside the 95 per cent region.

<u>Parametric Tests</u>

The method of parametric testing is to use the values obtained in the bivariate population of simulated scores to find maximum likelihood estimators in each population of the mean and variance of the Georgia Tech scores, the mean and variance of the opponent's scores, the covariance, and the coefficient of correlation. These values are used to calculate

$$u^* = \left\{\frac{1}{(1-\rho)^2}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]\right\}^{1/2}$$

Using u^*, the probability was found that a point (x,y) randomly drawn from a population having the parameters of the population of simulated scores would have a u such that $|u| \geq u^*$. The probabilities are shown in Table 10 along with the u^*'s. It may be seen from Table 10 that the number of actual scores which fell outside their 80 per cent confidence levels is five (games 3,5,6,11,12), four fell outside of 90 per cent confidence levels, and two (games 3 and 5) fell outside of 95 per cent confidence levels.

Table 10. A Display of the Statistics Used and the Results of the Parametric Tests

Game	u*	Pr(|u| > u*)
1	.734	.68
2	.133	.89
3	2.050	.04
4	.854	.40
5	3.042	.004
6	1.735	.09
7	1.015	.32
8	.107	.92
9	1.179	.24
10	.807	.42
11	1.362	.17
12	1.798	.07

NOTE: Critical value of u* at 80 per cent confidence level is u* > 1.28, at 90 per cent confidence level, u* > 1.65, and at 95 per cent, u* > 1.96.

CHAPTER IV

CONCLUSIONS AND RECOMMENDATIONS

The purpose of the football simulation model, its construction, and a test of its effectiveness have been described. This final chapter contains the conclusions which may be drawn from this thesis and recommendations for extending the results through further research.

Conclusions

From the statistical tests in the preceding chapter, it is reasonable to conclude that the simulation model does produce a population of scores from which the actual game score comes. The results of the nonparametric tests (three points out of 12 outside of 80 per cent confidence regions, two outside 90 per cent, and one outside of 95 per cent) may be ascribed to the level of the test. The parametric test results are not quite so favorable as this, but neither are they strongly contrary to the results of the nonparametric tests. The fact that the condition of bivariate normality is not met suggests that the parametric tests give no sound reason to doubt the favorable results of the nonparametric tests.

The fact that a population of scores can be produced which contains the actual game score is a significant first step into the scientific study of strategy selection in college football. Also of importance to further research is the small number of variable input parameters which are necessary for a simulation model. A simulation model would be of little final use to a football coach if he were then confronted with the problem of estimating accurately a large number of input parameters. A small number of input parameters eliminates the problem of accurate estimation by making it possible to assess many strategies under each of several input states and choose a minimax strategy to begin a game, modifying the strategy as the game progresses and more information is gained about the values of the input parameters.

That a scientific study of football is needed is demonstrated by the study of punting data. One of the most commonly held notions among football coaches, writers, and fans is that by punting on third down the kicking team has an advantage of a potential faked kick, thus forcing the receiving team to put more pressure on the kicker and ignore setting up a punt return. Because of this advantage, a third down punt is believed to give the receiving team the ball further from the goal line than a fourth down punt. However, the data which were gathered do not support this argument.

Regression lines passing through third down punting data are not significantly different than those passing through fourth down data. If anything, the data revealed a trend opposite to the prevailing belief. For both Georgia Tech and the opponents, third down punts tended to be less effective (though not significantly) than those kicked on fourth down. There may be sound arguments for third down punting, but this is not one of them.

The populations of simulated scores showed wide variances in their results. Much, and perhaps all, the variance may be attributed to the nature of the game of football. However, it is likely that some of the variance is caused by the decision-making procedure. A decision model which is a function of more of the game's relevant variables than the model one used in this study would produce more consistent decisions and, it would be reasonable to anticipate, tighter populations of simulated scores. The tighter populations would be of more value to a coach who was faced with using the populations to choose among several alternate strategies because they would be more sensitive to small differences between alternative strategies.

The objective of this study is to determine whether simulation can be an effective aid to a football coach in choosing among alternate offensive strategies. The conclusion is that it was effectively used in this study to assess the results of 12 selected strategies. The same methods which were effective in this study could be extended and, in places, developed in more detail, to give the coach a powerful method of analyzing his decisions.

Recommendations

Toward an attainable objective, the scientific evaluation of decision-making in college football, recommendations for further research will be made which remove some of the limitations of this paper, expand its scope, and adopt new procedures to the solution of the problem of how decisions should be made in this and related industries.

Many of the limitations of this study were caused by the size of the data sample. The process of gathering data from game descriptions and transferring them to punched cards is tedious, but once the data are gathered they are easily analyzed with the aid of digital computers. This study was based on data gathered from 12 games and from the problems which were encountered in analyzing some parameters, notably fumbles, penalties, and field goals, it may be concluded that an increase in data size by a factor of eight or ten would not be too large.

With these additional data the exact distributions of yardage gained on running and passing plays should be studied. They were assumed to be normally distributed in this study but intuitively a gamma distribution seems a better approximation. In the real-world game there are not so many occurrences of great losses of yardage as there are great gains and this is contrary to the assumptions of a normal distribution.

In expanding the scope of the research, the coach should be given more strategy options than the four which were allowed in this paper. For example, pass could be broken into short pass and long pass. Run could be broken into run into the line and run outside the line. These do represent legitimate strategic (rather than tactical) decision alternatives and should be treated as such.

The assumption that the offensive strategy decision is the only decision which affects the outcome of a play should be dropped in favor of a model which includes the defense's anticipations and preparations. This would allow the decision problem to be treated with the methods of game theory.

If the scope is expanded in these ways, the resulting model could serve as an effective training device for young players who need to gain experience in the complexities of both offensive and defensive strategy selection. They would be able to sit at a computer terminal, enter their decisions, and have them assessed by the computer and the results printed for them to make the next decisions.

An important topic which must eventually be dealt with before a thorough decision model can be constructed is the utilities associated with the possible game outcomes, winning, tying, and losing. The problem is best understood by considering the coach whose team has just scored a touchdown and who now trails the opponents by one point. Should he attempt a one- or a two-point conversion? The answer is inextricably tied to the utilities of the three possible outcomes. Involved in the utilities of the outcomes of the game may be: national football prestige, alumni relations, possible bowl bids bringing money to the school, and the morale of the team and its ability to function well the remainder of the season.

Some future research should be devoted to the topic of prediction of input parameters. The assumption of the simulations in this paper is that the input parameters are known. But a method of predicting them ahead of a game's actually being played must be developed.

APPENDIX

AN APPLICATION OF THE SIMULATION MODEL

A simple example of the type of problem to which the simulation model may be applied is described in this section.

In the Florida game of 1960, Georgia Tech's third down strategy (see Figure 11) was to run with probability .13, pass with probability .37, punt with probability .37, and attempt a field goal with probability .13. The simulated assessment of this strategy is shown in Figure 24.

To compare this strategy with one in which the coach did not kick on third down, the stragegy was changed to: run with probability .26 and pass with probability .74. The simulation was rerun and the results are shown in Figure 35 .

The results of the two simulations do not make it obvious which is the stronger strategy. Though the average score is much more favorable to Georgia Tech if they avoid kicking on third down, they only won one more of the simulated games (21 wins instead of 20 wins). To make a decision, the coach would have to specify the important criteria in choosing between the two populations of simulated scores.

NONPARAMETRIC TEST USED IN TESTING

THE RESULTS OF THE SIMULATION MODEL

n-1 samples are randomly drawn from the population Z of points in E^2. The problem is to determine whether another point, z_n, is a randomly chosen member of the population Z.

Each $z_i \in Z$, $i = 1,2,\ldots,n-1$, and z_n can be written as $z_i(x_i,y_i)$. The set of x_i's, $i = 1,2,\ldots,n$ are ranked from the lowest value to highest with ties broken by assigning each x_i which ties with one or more other x_j. The average value of the ranks which they tied for. After the x_i's are ranked, the set of y_i's are ranked in the same manner so that with each z_i, $i = 1,2,\ldots,n$, there may be associated the two ranks $r_x(z_i)$ and $r_y(z_i)$.

A z_i which is ranked in the middle on both the x and y rankings would have $r_x = \bar{r} = \frac{n+1}{2}$ and $r_y = \bar{r} = \frac{n+1}{2}$.

The probability,

$$P^* = \Pr\{\sqrt{(r_x-\bar{r})^2 + (r_y-\bar{r})^2} \geq \sqrt{(r_x(z_n)-\bar{r})^2 + (r_y(z_n)-\bar{r})^2}\}$$

is the probability of choosing a z_i at random such that the Euclidean distance of the point in the plane representing its ranks, $(r_x(z_i), r_y(z_i))$, is as far or farther from the point $(\bar{r},\bar{r})$ than is the point $(r_x(z_n), r_y(z_n))$. If P* is small, the null hypothesis, $z_n \in Z$ will be rejected.

Table 11. A Display of the Numerical Data Used in the Regression Analysis of Georgia Tech and Opponents' Punting

Team	Down	Data Group	Line	$\sum x_i$	$\sum y_i$	$\sum x_i^2$	$\sum y_i^2$	$\sum x_i y_i$	n
Georgia Tech	1-3	all data	L_1	344	920	9,352	56,128	21,614	17
	4	all data	L_2	1821	3402	85,151		144,972	45
	4	0-40 yard line	L_3	626	1516	18,784	102,308	42,448	23
	4	40+ yard line excluding punts over goal line	L_4	851	1406	46,227	124,476	75,004	17
	1-4	0-40 yard line	L_9	970	2436	28,136	158,436	64,062	40
Opponents	1-3	all data	L_5	382	1099	9,562	62,941	21,073	23
	4	all data	L_6	1916	3611	77,287		134,331	56
	4	0-40 yard line	L_7	908	2027	26,008	127,449	53,964	36
	4	40+ yard line excluding punts over goal line	L_8	667	1109	32,183	89,307	53,213	14
	1-4	0-40 yard line	L_{10}	1290	3126	35,570	190,390	75,037	59

Table 11. A Display of the Numerical Data Used in the Regression Analysis of Georgia Tech and Opponents' Punting (Continued)

Line	$\bar{x}$	$\bar{y}$	$\hat{\beta} = \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sum(x_i-\bar{x})^2}$	$\hat{\alpha} = \bar{y}-\hat{\beta}\bar{x}$	$\hat{s}^2 = \frac{\sum[y_i-(\hat{\beta}x_i+\alpha)]^2}{n-2}$
L_1	20.24	54.12	1.25	28.8	176.3
L_2	40.46	75.60	.64	49.9	
L_3	27.27	65.91	.68	47.4	75.2
L_4	50.59	82.70	1.08	34.7	162.5
L_9	24.25	60.90	1.08	34.7	125.7
L_5	16.68	47.78	.87	33.3	387.7
L_6	30.64	64.48	.92	33.0	
L_7	25.22	56.31	.91	33.2	315.4
L_8	47.64	79.21	1.20	21.0	96.0
L_{10}	21.86	52.98	.89	33.5	334.7

Test 1. Test of the Hypothesis that the Variances about the Four Punting Regression Lines Are Equal

---

<u>H_0:</u> $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 = \sigma^2$

<u>H_A:</u> At least one $\sigma_i^2 \neq \sigma^2$

<u>Test Statistic:</u>
(Bartlett's Test)

$$z = -\frac{\sum_{i=1}^{4} f_i \ \ell n \left(\frac{s_i^2}{s^2}\right)}{1 + 3\,\frac{1}{(4-1)} \sum_{i=1}^{4} \left(\frac{1}{f_i} - \frac{1}{f}\right)}$$

where

$$s^2 = \frac{\sum_{i=1}^{4} f_i \ s_i^2}{f}, \quad f = \sum_{i=1}^{k} f_i, \quad f_i = \text{number of degrees of freedom in the ith sample variance}$$

<u>Critical Region:</u> $z > \chi^2(k, .975) = \chi^2(3, .975) = 9.35$

<u>Test:</u>

$$z = \frac{-\left[15 \ \ell n\left(\frac{176.3}{253.7}\right) + 21 \ \ell n\left(\frac{387.7}{253.7}\right) + 34 \ \ell n\left(\frac{315.4}{253.7}\right) + 21 \ \ell n\left(\frac{75.2}{253.7}\right)\right]}{1 + \frac{1}{9}\left[\frac{1}{15} + \frac{1}{21} + \frac{1}{34} - \frac{4}{91}\right]}$$

$$= \frac{14.797}{1.016}$$

$$= 14.558$$

<u>Conclusion:</u> Reject H_0.

Test 2. Test of the Hypothesis that the Variances about the Two Regression Lines for Georgia Tech Punting Are Equal

---

H_0: $\sigma_1^2 = \sigma_3^2$

H_A: $\sigma_1^2 \neq \sigma_3^2$

Test Statistic: $z = \dfrac{s_1^2}{s_2^2}$

Critical Region: $z > F(n_1-2,\ n_3-2;\ .975) = F(17,\ 21;\ .975) = 2.48$

Test: $z = \dfrac{176.3}{75.2}$

$= 2.32$

Conclusion: Do not reject H_0.

Test 3. Test of the Hypothesis that the Slopes of the Two Punting Regression Lines for Georgia Tech are Equal

---

H_0: $\beta_1 = \beta_3$

H_A: $\beta_1 \neq \beta_3$

Test Statistic:

$$z = \frac{\hat{\beta}_1 - \hat{\beta}_3}{s\sqrt{\dfrac{1}{\sum_{i=1}^{n_1} (x_{1i}-\bar{x}_1)^2} + \dfrac{1}{\sum_{i=1}^{n_3} (x_{3i}-\bar{x}_3)^2}}}$$

where $$s^2 = \frac{(n_1-2)s_1^2 + (n_3-2)s_3^2}{n_1 + n_3 - 4}$$

Critical Region: $|z| > t(n_1+n_3-4;\ .975) = t(36,\ .975) = 2.031$

Test:

$$s^2 = \frac{(15)(176.3) + (21)(75.2)}{36} = 117.33$$

$$z = \frac{1.25 - .68}{10.8\sqrt{\dfrac{1}{2391.3} + \dfrac{1}{1746.6}}} = \frac{.57}{.33}$$

$$= 1.727$$

Conclusion: Do not reject H_0.

Test 4. Test of the Hypothesis that the Two Regression Lines for Georgia Tech Punting are Concurrent

---

$\underline{H_0:}$ $L_1 = L_3$

$\underline{H_A:}$ $L_1 \neq L_3$

Test Statistic:
$$z = \frac{(\bar{y}_3-\bar{y}_1) - \beta_{13}(\bar{x}_3-\bar{x}_1)}{s \quad \frac{1}{n_1} + \frac{1}{n_3} + (\bar{x}_3-\bar{x}_1)^2 / \sum\limits_{i=1,3} \sum\limits_{v=1}^{n_1} (x_{iv}-\bar{x}_i)^2}$$

where
$$s^2 = \frac{\sum\limits_{i=1,3} \sum\limits_{v=1}^{n_1} y_{iv}^2 - \sum\limits_{i=1,3} \frac{\left(\sum\limits_{v=1}^{n_1} y_{iv}\right)^2}{n_i} - \frac{\left[\sum\limits_{i=1,3} \sum\limits_{v=1}^{n_i} y_{iv}(x_{iv}-\bar{x}_i)\right]^2}{\sum\limits_{i=1,3} \sum\limits_{v=1}^{n_i} (x_{iv}-\bar{x}_i)^2}}{n_1 + n_3 - 3}$$

Critical Region: $|z| > t(n_1+n_3-3;\ .975) = t(37;\ .975) = 2.031$

Test: $s^2 = 124.1$

$$z = \frac{4.20}{11.14\sqrt{\frac{1}{17} + \frac{1}{23} + \frac{(7.03)^2}{4068}}}$$

$$= 1.12$$

Conclusion: Do not reject H_0.

Test 5. Test of the Hypothesis that the Variances about the Two Regression Lines for Opponents' Punting Are Equal

---

<u>H_0:</u> $\sigma_t^2 = \sigma_7^2$

<u>H_A:</u> $\sigma_5^2 \neq \sigma_7^2$

<u>Test Statistic:</u> $z = \dfrac{s_5^2}{s_7^2}$

<u>Critical Region:</u> $z > F(n_5-2,\ n_7-2;\ .975) = F(21,\ 33;\ .975) = 2.28$

<u>Test:</u> $z = \dfrac{387.7}{315.4} = 1.23$

<u>Conclusion:</u> Do not reject H_0.

Test 6. Test of the Hypothesis that the Slopes of the Two Regression Lines for Opponents' Punting are Equal

---

<u>H_0:</u> $\beta_5 = \beta_7$

<u>H_A:</u> $\beta_5 \neq \beta_7$

<u>Test Statistic:</u> See Test 3.

<u>Critical Region:</u> $|z| > t(n_5+n_7-4;\ .95) = t(51;\ .95) = 1.68$

<u>Test:</u> $s^2 = 370$

$$z = \frac{1.12 - .91}{19.2\sqrt{\dfrac{1}{3517} + \dfrac{1}{11{,}732}}}$$

$$= .569$$

<u>Conclusion:</u> Do not reject H_0.

Test 7. Test of the Hypothesis that the Two Regression Lines for Opponents' Punting Are Concurrent

---

$\underline{H_0}$: $L_5 = L_7$

$\underline{H_A}$: $L_5 \neq L_7$

Test Statistic: See Test 4.

Critical Region: $|z| > t(n_5+n_7-3;\ .975) = t(56;\ .975) = 2.005$

Test: $s^2 = 387.5$

$$z = \frac{-\ 8.50 - (.89)(-8.54)}{19.68\sqrt{\frac{1}{23} + \frac{1}{36} + \frac{72.93}{15,249}}}$$

$$= -\ .165$$

Conclusion: Do not reject H_o.

Test 8. Test of the Hypothesis that the Variances in the Populations of Kickoffs by Georgia Tech and by Opponents Are Equal

---

$\underline{H_0}$: $\sigma_1^2 = \sigma_2^2$

$\underline{H_A}$: $\sigma_1^2 \neq \sigma_2^2$

Test Statistic: $z = \dfrac{s_1^2}{s_2^2}$

Critical Region: $z > F(n_1, n_2; .975) = F(25, 44; .975) = 1.98$

Test: $z = \dfrac{6863/25}{2734/44} = 4.42$

Conclusion: Reject H_0.

Test 9. Test of the Hypothesis that the Means of the Populations of Kickoffs by Georgia Tech and by Opponents Are Equal

---

H0: $\mu_1 = \mu_2$

HA: $\mu_1 \neq \mu_2$

Test Statistic: (Behrens-Fisher Test)

$$z = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$$

with f degrees of freedom where

$$f = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{(s_1^2/n_1)^2/(n_1-1) + (s_2^2/n_2)^2/(n_2-1)}$$

Critical Region: $|z| > t(f;\ .975) = t(31.7;\ .975) = 2.042$

Test:

$$f = \frac{(10.56 + 1.38)^2}{\frac{(10.56)^2}{25} + \frac{(1.38)^2}{44}} = 31.7$$

$$z = \frac{29.35 - 27.69}{\sqrt{\frac{6863}{(26)(25)} + \frac{2734}{(45)(44)}}} = .481$$

Conclusion: Do not reject H_o.

Test 10. Test of the Hypothesis that the Probability of a Pass Interception by Georgia Tech is Equal to the Probability of a Pass Interception by Opponents

---

H_0: $\theta_1 = \theta_2 = \theta$

H_A: $\theta_1 \neq \theta_2$

Test Statistic:

$$z = \frac{(\theta_1 - \frac{1}{2\,n_1}) - (\theta_2 - \frac{1}{2\,n_2})}{\left[(\theta)(1-\theta)(\frac{1}{n_1} + \frac{1}{n_2})\right]^{1/2}}$$

Critical Region: $z > N^{-1}(.975) = 1.96$

Test:

$$z = \frac{(.058 - \frac{1}{482}) - (.031 - \frac{1}{508})}{\left[(.044)(.956)(\frac{1}{241} + \frac{1}{254})\right]^{1/2}}$$

$$= 1.46$$

Conclusion: Do not reject H_0.

Test 11. Test of the Hypothesis that the Variance about 463 Population Means is Equal to the Variance about 66 Population Means

---

<u>H_0:</u> $\sigma_1^2 = \sigma_2^2$

<u>H_A:</u> $\sigma_1^2 \neq \sigma_2^2$

<u>Test Statistic:</u> $z = \dfrac{(s_1^2 - s_2^2)/(f_1 - f_2)}{s_2^2/f_2}$

<u>Critical Region:</u> $z > F(f_1 - f_2, f_2; .975) = F(397, 650; .975) \approx 1.15$

<u>Test:</u> $z = \dfrac{13{,}056.30/397}{19{,}628.90/(1113-463)} = 1.09$

<u>Conclusion:</u> Do not reject H_0.

```
IDENTIFICATION DIVISION.
PROGRAM-ID.   FOOTBALL SIMULATION.
ENVIRONMENT DIVISION.
CONFIGURATION SECTION.
SOURCE-COMPUTER. B-5500.
OBJECT-COMPUTER. B-5500.
INPUT-OUTPUT SECTION.
FILE-CONTROL.
     SELECT IP-FILE ASSIGN TO READER.
     SELECT OP-FILE ASSIGN TO PRINTER, TAPE.
DATA DIVISION.
FILE SECTION.
FD   IP-FILE
     LABEL RECORDS ARE STANDARD VALUE ID IS "CARD    " DATA
     RECORD IS IP-REC.
01   IP-REC.
     03  INNER.
         05 HARAMETERS-IP            PICTURE IS 9(14).
         05 FILLER                   PICTURE IS X(66).
     03  INNER-1 REDEFINES INNER.
         05 INDEX                    PICTURE IS  9.
         05 HDING                    PICTURE IS X(40).
         05 FILLER                   PICTURE IS X(39).
FD   OP-FILE
     LABEL RECORDS ARE STANDARD VALUE ID IS "PRINTER" DATA
     RECORD IS VEC.
01   VEC.
     03  VECA.
         04  FILLER                  PICTURE IS X(10).
         04  GP                      PICTURE IS X(20).
         04  GP1                     PICTURE IS      9.
         04  GP2                     PICTURE IS X(32).
         04  GP2A REDEFINES GP2.
             05  PMU                 PICTURE IS   +++.99.
             05  FILLER              PICTURE IS X(7).
             05  PSD                 PICTURE IS     ZZ.99.
             05  FILLER              PICTURE IS X(14).
     04  GP2B REDEFINES GP2A.
         05 D-PRA OCCURS 4 TIMES.
             06 FILLER               PICTURE IS    XX.
             06 D-PR                 PICTURE IS  ZB99.
             06 FILLER               PICTURE IS    XX.
         04  GP3                     PICTURE IS Z.999.
         04  FILLER                  PICTURE IS X(42).
     03  D-VEC REDEFINES VECA.
         04  FILLER                  PICTURE IS X(11).
         04  D-Q                     PICTURE IS      9.
         04  FILLER                  PICTURE IS    XX.
         04  TIME-D.
             06 MIN-D                PICTURE IS    ZZ.
             06 PUNC                 PICTURE IS      X.
             06 SEC-D                PICTURE IS    99.
         04  FILLER                  PICTURE IS   XXX.
         04  T-D                     PICTURE IS      9.
         04  FILLER                  PICTURE IS  XXXX.
         04  D-D                     PICTURE IS      9.
         04  FILLER                  PICTURE IS    XX.
         04  X-D                     PICTURE IS   ZZ9.
         04  FIRST-D                 PICTURE IS  Z(5).
         04  FILLER                  PICTURE IS    XX.
         04  DCSN                    PICTURE IS    XX.
         04  FILLER                  PICTURE IS   XXX.
```

Figure 10. Listing of the COBOL Simulation Model

```
               04  GAIN.
                   06  LOSS              PICTURE IS     X.
                   06  ABS-O             PICTURE IS    Z9.
               04  PEN-O.
                       06  FILLER        PICTURE IS   XX.
                       06  AG-O          PICTURE IS     9.
                   06  FILLER            PICTURE IS     X.
                       06  YG-O          PICTURE IS    Z9.
                   06  AC-O              PICTURE IS   ZZZ.
               04  FILLER                PICTURE IS    XX.
               04  COND                  PICTURE IS   XXX.
               04  SCO-O OCCURS 2 TIMES PICTURE IS ZZZ9.
               04  CHNG-POS              PICTURE IS  ZZZZ.
                   04  FILLER            PICTURE IS X(03).
                   04  EOQ               PICTURE IS X(03).
               04  TRACER                PICTURE IS XXXXX.
               04  FRNAER                PICTURE IS Z(8).
               04  FILLER                PICTURE IS X(8).
           04  BRACER PICTURE IS +++++.999.
           03  H-VEC REDEFINES O-VEC.
               05  H-VEC-B               PICTURE IS X(77).
               05  H-VEC-C               PICTURE IS X(33).
           03  SUMMARY-WRITER REDEFINES H-VEC.
           05  FILLER                    PICTURE IS X(30).
           05  OPPOSITION                PICTURE IS X(20).
           05  FIN-SCO OCCURS 2 TIMES    PICTURE IS ZZZZ9.
           05  FILLER                    PICTURE IS X(50).
           03  GRAF REDEFINES SUMMARY-WRITER.
               05  FILLER      PC X(28).
               05  SKD         PC ZZ.
               05  GR-ROW              PC X(52).
           03  STUFF REDEFINES GRAF.
               04 FILLER       PC X(10).
               04 ST1          PC X(19).
               04 ST2.
                   05  ST3 OCCURS 27 TIMES PC Z79.
           03  FILLER                    PICTURE IS X(10).
       WORKING-STORAGE SECTION.
       77  M                             PICTURE IS 9(8).
       77  D                             PICTURE IS    9.
       77  X                             PICTURE IS  S999.
       77  PINV                          PICTURE IS S9V999.
       77  T                             PICTURE IS    99.
       77  F                             PICTURE IS   S99.
       77  PR                            PICTURE IS   999.
       77  SW1                           PICTURE IS     9.
       77  SIG                           PICTURE IS   999V999.
       77  MU                            PICTURE IS  S999V999.
       77  SEC-HALF                      PICTURE IS     9.
       77  XA                            PICTURE IS   999.
       77  XB                            PICTURE IS   999.
       77  VAL                           PICTURE IS S999.
       77  P-HOLD                        PICTURE IS S999.
       77  DEC                           PICTURE IS     9.
       77  Q                             PICTURE IS     9.
       77  TIM                           PICTURE IS    99.
       77  OC                            PICTURE IS     9.
       77  TP                            PICTURE IS     9.
       77  XP                            PICTURE IS    99.
       77  PROB                          PICTURE IS  V99 .
       77  HD                            PICTURE IS 99V999.
       77  PSTN                          PICTURE IS     9.
```

Figure 10. Listing of the COBOL Simulation Model (Continued)

```
       77  H                                    PICTURE IS   999.
       77  H1                                   PICTURE IS   999.
       77  HAL                                  PICTURE IS  S999.
       77  BUM                                  PICTURE IS 9(8).
       77  SIGPINV                              PICTURE IS S999V999.
       77  SLOPE                                PICTURE IS 99V99.
       77  Y-INT                                PICTURE IS 99V99.
       77  REP  PICTURE IS 99.
       77  SIMU  PICTURE IS 99 VALUE IS 01.
       77  RN-MULT               PC 9(9).
       77  RX  PC 9.
       77  R          PC 99.
       77  N  PC 99 VA 25.
       77  RA         PC 99.
       77  FACT       PC 9999V99.
       77  SIGMA      PC S999V9999.
       77  XSQ        PC 9(8)V99.
       77  YSQ        PC 9(8)V99.
       77  DSQ        PC 9(8)V99.
       77  K        PC 99999 VA 31621.
       77  RHO        PC S9V999.
       77  COVAR PC S9999999V99.
       77  CORR PC SV999.
       77  XY PC 9999999.
       77  DIMU PC 99 VA 01.
       01  TIES.
           03  TIE OCCURS 2 TIMES PICTURE IS S999V99.
       01  RANDOM-NUMBER PC 9(18).
       01  TANDEM-NUMBER REDEFINES RANDOM-NUMBER.
           04  FILLER            PC 9(9).
           04  MOD-OP            PC 9(9) VA 173964213.
           04  RN REDEFINES MOD-OP.
               06  RNA.
                   08  RND     PC 99.
                   08  RNB     PC  9.
               06  RNC           PC 9(6).
       01  SCORE.
           02  SCD  OCCURS 2 TIMES       PICTURE IS   999.
       01  INPUT-PARAMETERS.
           03 TMP          OCCURS 2 TIMES.
               08  PLAA OCCURS 4 TIMES.
                   09  PARA OCCURS 2 TIMES PICTURE IS 999V999.
       01  DECISION-PARAMETERS.
           03 D-TEAM OCCURS 2 TIMES.
               05 D-DOWN OCCURS 4 TIMES.
                   07  D-PLAY OCCURS 4 TIMES PICTURE IS 999.
       01  TIME.
           02 MIN     PICTURE IS 99 VA 15.
           02 SEC     PICTURE IS 99 VA 00.
       01  ADAC                                 PICTURE IS    99.
       01  BDBC REDEFINES ADAC.
           03  BC                               PICTURE IS     9.
           03  BD                               PICTURE IS     9.
       01  PARAMETERS-IP.
           04  TM                       PICTURE IS     9.
           04  PTYPE                    PICTURE IS     9.
           04  POSTN                    PICTURE IS     9.
           04  P-VAL                    PICTURE IS 99V999.
           04  FILLER                   PICTURE IS  X(6).
       01  DCSN-IP REDEFINES PARAMETERS-IP.
           04 D-TM                      PICTURE IS     9.
           04 D-DN                      PICTURE IS     9.
```

Figure 10. Listing of the COBOL Simulation Model (Continued)

```
           04 D-PL OCCURS 4 TIMES.
                06 D-P                    PICTURE IS   999.
       01  PARAMETER-TABLE.
           03 PA OCCURS 350 TIMES PICTURE IS 9(14).
       01  GAME-HEADERS.
           03   G-HEADING    OCCURS 12 TIMES.
                05 G-SCORE   OCCURS   2 TIMES   PICTURE  99.
                05 G-GAM                        PICTURE  99.
                05 FILLER                       PICTURE X(12).
                05 G-TEAM                       PICTURE X(20).
       01  SIMULATED-SCORES.
           03   S-GAME   OCCURS   2 TIMES.
                05    S-SCO OCCURS 30 TIMES.
                      07   S-TEAM OCCURS 2 TIMES PICTURE 99.
                      07   S-RANK OCCURS 2 TIMES PICTURE 99V99.
                05    S-AVG OCCURS 2 TIMES       PC 9999V99.
                05    S-VAR OCCURS 2 TIMES        PICTURE 999999V99.
       01   GRAPH.
            03   Y-CORD   OCCURS 26 TIMES.
                 05   X-CORD   OCCURS 26 TIMES PC XX.
       01   GRAPH-SYM.
            03   BORK             PC XX.
            03   CORK             PC X.
            03   EORK             PC X.
       CONSTANT SECTION.
       01   NORMAL-TABLE.
            02   N0   PICTURE 9V999 USAGE COMPUTATIONAL VALUE  0.000.
            02   N1   PICTURE 9V999 USAGE COMPUTATIONAL VALUE  0.025.
            02   N2   PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.050.
            02   N3   PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.075.
            02   N4   PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.100.
            02   N5   PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.125.
            02   N6   PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.151.
            02   N7   PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.176
            02   N8   PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.202.
            02   N9   PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.227.
            02   N10 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.253.
            02   N11 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.278.
            02   N12 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.305.
            02   N13 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.332.
            02   N14 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.358.
            02   N15 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.385.
            02   N16 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.412.
            02   N17 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.440.
            02   N18 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.468.
            02   N19 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.496.
            02   N20 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.525.
            02   N21 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.554.
            02   N22 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.583.
            02   N23 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.613.
            02   N24 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.643.
            02   N25 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.674.
            02   N26 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.706.
            02   N27 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.739.
            02   N28 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.772.
            02   N29 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.806.
            02   N30 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.842.
            02   N31 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.878.
            02   N32 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.915.
            02   N33 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.954.
            02   N34 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 0.995.
            02   N35 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 1.037.
```

Figure 10. Listing of the COBOL Simulation Model (Continued)

```
        02  N36 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 1.080.
        02  N37 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 1.126.
        02  N38 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 1.175.
        02  N39 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 1.226.
        02  N40 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 1.281.
        02  N41 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 1.340.
        02  N42 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 1.405.
        02  N43 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 1.476.
        02  N44 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 1.555.
        02  N45 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 1.645.
        02  N46 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 1.750.
        02  N47 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 1.881.
        02  N48 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 2.054.
        02  N49 PICTURE 9V999 USAGE COMPUTATIONAL VALUE 2.330.
    01 NORMAL-TABLE-A REDEFINES NORMAL-TABLE.
        02 PHI-INV PICTURE 9V999 USAGE IS COMPUTATIONAL
            OCCURS 50 TIMES.
    PROCEDURE DIVISION.
    L1.
        MOVE SPACES TO GRAPH.
        OPEN INPUT IP-FILE OUTPUT OP-FILE.
    HEADING-READER.
        ADD 1 TO H.
        READ IP-FILE RECORD AT END  CLOSE IP-FILE WITH RELEASE.
        MOVE HDING TO G-HEADING(H).
        IF H≠12 GO TO HEADING-READER ELSE MOVE 0 TO H.
    READ-IN.
        ADD 1 TO H.
        READ IP-FILE RECORD AT END SUBTRACT 1 FROM H CLOSE IP-FILE
            WITH RELEASE MOVE 1 TO REP GO TO L4B.
        MOVE HARAMETERS-IP TO PA(H).
        GO TO READ-IN.
    L4.
        ADD 1 TO REP.
        IF REP =N+1 PERFORM SUMMARIZATION THRU SUM-2.
        IF REP=N+1 MOVE 1 TO REP ADD 1 TO DIMU ELSE
        MOVE 1 TO DIMU GO TO FOOTBALL-GAME.
        IF RX = 1 AND DIMU=13 GO TO NULL.
        IF DIMU=13 MOVE 1 TO DIMU MOVE 0 TO H1 MOVE 31627 TO K
            MOVE 1 TO RX.
    L4B.
        PERFORM PSEUDO-READ.
        IF TM=3 GO TO L6.
        MOVE P-VAL TO PARA(TM,PTYPE,POSTN).
        GO TO L4B.
    L5.
        MOVE SPACES TO VEC.
        WRITE VEC BEFORE ADVANCING 6 LINES.
        MOVE "    GAME PARAMETERS" TO GP WRITE VEC BEFORE ADVANCING
            2 LINES.
        PERFORM B1          VARYING T FROM 1 BY 1 UNTIL T=2.
    B1.
        IF T=1 MOVE "GEORGIA TECH        " TO GP ELSE MOVE G-TEAM
            IN G-HEADING(SIMU) TO GP.  WRITE VEC PERFORM ADV.
        MOVE "          PR(INCOMPLETE PASS)   " TO GP2.
        MOVE PARA(T,3,1) TO GP3.
        WRITE VEC PERFORM ADV.
        MOVE "          PR(PASS NOT THROWN)   " TO GP2.
        MOVE PARA(T,3,2) TO GP3.
        WRITE VEC PERFORM ADV.
        MOVE "MEAN   STANDARD DEVIATION      " TO GP2.
        WRITE VEC MOVE SPACES TO VEC.
```

Figure 10. Listing of the COBOL Simulation Model (Continued)

```
       MOVE "          RUN       " TO GP.  MOVE PARA(T,1,1) TO PMU.
            MOVE PARA(T,1,2) TO PSD. WRITE VEC. MOVE SPACES TO VEC.
       MOVE "          PASS POS" TO GP.  MOVE PARA(T,2,1) TO PMU.
            MOVE PARA(T,2,2) TO PSD. WRITE VEC.
       MOVE "          PASS NEG" TO GP.
       SUBTRACT PARA(T,4,1) FROM  0 GIVING PMU.
            MOVE PARA(T,4,2) TO PSD. WRITE VEC.
       PERFORM ADV.
       PERFORM ADV.
   L6.
       PERFORM PSEUDO-READ.
       IF D-TM=3 GO TO FOOTBALL-GAME.
       PERFORM L6A VARYING X FROM 1 BY 1 UNTIL X=4.
   L6A.
       MOVE D-P IN D-PL(X) TO D-PLAY(D-TM,D-DN,X).
   L6B.
       GO TO L6.
   L7.
       MOVE " DECISION PARAMETERS" TO GP WRITE VEC.
       PERFORM ADV 2 TIMES.
       PERFORM L7A THRU L7D VARYING T FROM 1 BY 1 UNTIL T=2.
   L7A.
       IF T=1 MOVE "GEORGIA TECH          " TO GP ELSE MOVE G-TEAM
            IN G-HEADING(SIMU) TO GP.  WRITE VEC PERFORM ADV.
       MOVE "   PERCENT OF PLAYS THAT ARE:    " TO GP2 WRITE VEC.
       MOVE "    RUN     PASS     PUNT       FG   " TO GP2.
       WRITE VEC PERFORM ADV.
       PERFORM L7B THRU L7D VARYING D FROM 1 BY 1 UNTIL D=4.
   L7B.
       MOVE "              DOWN " TO GP. MOVE D TO GP1.
       PERFORM L7C VARYING X FROM 1 BY 1 UNTIL X=4.
   L7C.
       MOVE D-PLAY(T,D,X) TO D-PR IN D-PRA(X).
   L7D.
       WRITE VEC. MOVE SPACES TO VEC.
       IF D=4 PERFORM ADV.
   FOOTBALL-GAME.
       MOVE 1 TO Q.
       MOVE ZEROES TO SCO(1).
       MOVE ZEROES TO SCO(2).
       PERFORM HDR1.
       PERFORM COIN-TOSS.
       PERFORM K-OFF.
       MOVE 1 TO D.
   L8.
       MOVE SPACES TO VEC.
   DECISION.
       PERFORM RNG.
       IF RND < D-PLAY(T,D,1) MOVE 1 TO DEC GO TO RUNN.
       IF RND < D-PLAY(T,D,1) + D-PLAY(T,D,2) MOVE 2 TO DEC GO TO
            PASS.
       IF RND < D-PLAY(T,D,1) + D-PLAY(T,D,2) + D-PLAY(T,D,3) MOVE
            3 TO DEC GO TO KICK.
       MOVE 4 TO DEC GO TO FIELD-GOAL.
   RUNN.
       MOVE 32 TO TIM PERFORM TIMER.
       PERFORM RNG.
       IF RNA < 028 GO TO M1.
       MOVE PARA(T,1,1) TO MU MOVE PARA(T,1,2) TO SIG.
       PERFORM COM. PERFORM GAINER.  GO TO M6.
   PASS.
       MOVE 25   TO TIM PERFORM TIMER.
```

Figure 10. Listing of the COBOL Simulation Model (Continued)

```
           PERFORM RNG.
           IF RNA < 028 GO TO M1.
           PERFORM RNG.
           IF RND < PARA(T,3,2)*100 SUBTRACT PARA(T,4,1) FROM 0 GIVING
               MU MOVE PARA(T,4,2) TO STG MOVE 3 TO
               TIM PERFORM TIMER GO TO M2 ELSE MOVE PARA(T,2,1) TO MU
               MOVE PARA(T,2,2) TO STG.
           PERFORM RNG.
           IF RND < PARA(T,3,1)*100  ADD 29 TO SEC
               MOVE ZEROES TO VAL GO TO M6.
           PERFORM RNG.
           IF RNA<046  MOVE 11.9 TO MU MOVE 15.524 TO
               STG PERFORM COM GO TO M7.
       M2.
           PERFORM COM.
           PERFORM GAINER.
       M6.
           PERFORM RNG.
           IF RNA < 039 GO TO PENALTY.
       M3.
           ADD VAL TO X.
           IF X> 99 GO TO TOUCHDOWN.
           IF X <   0 GO TO SAFETY.
           SUBTRACT VAL FROM F.
           IF F < 0 MOVE 1 TO D MOVE 10 TO F  GO TO EOQ-TEST.
           IF D = 4 PERFORM CPOS  GO TO EOQ-TEST.
           ADD 1 TO D GO TO EOQ-TEST.
       M1.
           SUBTRACT .587 FROM 0 GIVING MU.
           MOVE 5.831 TO STG.
           PERFORM COM.
           PERFORM RNG.
           IF RND < 50 GO TO M3.
       M4.
           ADD VAL TO X.
       M5.
           MOVE 1 TO D MOVE 10 TO F.
           IF T=1 MOVE 2 TO T ELSE MOVE 1 TO T.
           SUBTRACT X FROM 100 GIVING X.
           IF X> 99 GO TO TOUCHDOWN.
           IF X ≤ 0  MOVE 20 TO X.
           GO TO EOQ-TEST.
       M7.
           ADD VAL TO X IF X> 99 MOVE 80 TO X GO TO M5.
       TOUCHDOWN.
           PERFORM RNG. IF RNA < 300 ADD
               6 TO SCO(T) GO TO TD-1. IF RNA < 954 ADD 7 TO
               SCO(T) ELSE ADD 8 TO SCO(T).
       TD-1.
           MOVE 5 TO TIM PERFORM TIMER.
           IF T=1 MOVE 2 TO T ELSE MOVE 1 TO T.
           MOVE 1 TO D  MOVE 10 TO F.
           PERFORM K-OFF.
       TD-2.
           GO TO EOQ-TEST.
       SAFETY.
           IF T=2 ADD 2 TO SCO(1) ELSE
               ADD 2 TO SCO(2).
           PERFORM TD-1.
           ADD 20 TO X.
           GO TO EOQ-TEST.
       KICK.
```

Figure 10. Listing of the COBOL Simulation Model (Continued)

```
           IF X > 65  GO TO FIELD-GOAL.
           MOVE 30 TO TIM PERFORM TIMER. PERFORM RNG.
           IF RND > 01 GO TO K2 ELSE  SUBTRACT 10
               FROM X. IF RND=00 AND D≠4 ADD 10 TO F.
           IF ABS-D≠0 AND X<0 GO TO SAFETY.
           IF ABS-D≠0 GO TO EOQ-TEST.
           IF X<0 PERFORM C-POS GO TO TOUCHDOWN.
       K1.
           PERFORM C-POS.  GO TO EOQ-TEST.
       K2.
           IF T=1 AND X≤40 MOVE 1.08 TO SLOPE MOVE 34.70 TO Y-INT
               MOVE 11.21 TO STG GO TO K2A.
           IF T=1 MOVE 1.08 TO SLOPE MOVE 34.70 TO Y-INT MOVE 12.74 TO
               STG GO TO K2A.
           IF T=2 AND X≤40 MOVE 0.89 TO SLOPE MOVE 33.50 TO Y-INT
               MOVE 18.39 TO STG GO TO K2A.
           MOVE 1.20 TO SLOPE MOVE 21.00 TO Y-INT MOVE 9.80 TO SIG.
       K2A.
           MULTIPLY SLOPE BY X GIVING MU ADD Y-INT TO MU.
           PERFORM COM.
           IF VAL≥100 MOVE 80 TO VAL.
           COMPUTE P-HOLD=VAL - X.
           PERFORM RNG. IF RND≥03 GO TO K3.
               ADD 15
               TO X MOVE 10 TO F MOVE 1 TO D GO TO EOQ-TEST.
       K3.
           IF RND ≥ 04 GO TO  K4.
               IF F>5 AND P-HOLD ≥ MU - X
               + 5 ADD P-HOLD TO X PERFORM
               C-POS MOVE 10 TO F MOVE 1 TO D GO TO EOQ-TEST.
           ADD 5 TO X. SUBTRACT 5 FROM F. IF F≤0 MOVE 10
               TO F MOVE 1 TO D. GO TO EOQ-TEST.
       K4.
           IF RND ≥ 06 GO TO K5.
           IF X <10 DIVIDE 2 INTO X GIVING HAL ROUNDED ELSE MOVE 5 TO
               HAL.
           IF MU - X         + HAL≥ P-HOLD  ADD P-HOLD TO
               X PERFORM C-POS  GO TO EOQ-TEST.
           SUBTRACT HAL FROM X ADD HAL TO F.
               GO TO EOQ-TEST.
       K5.
           IF RND ≥ 10 GO TO K6.
           ADD P-HOLD TO X MOVE 10 TO F MOVE 1 TO D
               GO TO EOQ-TEST.
       K6.
           ADD P-HOLD TO X.  IF X≤0 GO TO TOUCHDOWN.
           PERFORM C-POS. GO TO EOQ-TEST.
       FIELD-GOAL.
           IF X<55 GO TO KICK.
           SUBTRACT X FROM 100 GIVING VAL.  MOVE VAL TO ABS-D.
           PERFORM FG.
           MOVE 5 TO TIM PERFORM TIMER.
           PERFORM CPOS.
           IF SW1=1 PERFORM K-OFF ELSE MOVE 20 TO X.
           MOVE 0 TO SW1.
           IF EOQ="EOQ" GO TO Q-END.
           GO TO L8.
       PENALTY.
           PERFORM RNG.
           MOVE VAL TO P-HOLD.
           IF RNA < 629 GO TO PEN-1.
           PERFORM RNG.
```

Figure 10. Listing of the COBOL Simulation Model (Continued)

```
           IF RND < 40 MOVE 5 TO VAL GO TO PEN-2.
           IF RND < 75 MOVE 15 TO VAL  GO TO PEN-2.
           PERFORM RNG.
           DIVIDE 3 INTO RND GIVING VAL ROUNDED.
               GO TO PEN-2.
       PEN-1.
           PERFORM RNG.
           IF RNA < 625 MOVE 5 TO VAL  GO TO PEN-3.
           IF RNA < 875 PERFORM RNG DIVIDE 5 INTO RND GIVING VAL ROUNDED
               GO TO PEN-3.
           MOVE 15 TO VAL  GO TO PEN-3.
       PEN-2.
           IF VAL + X ≥ 100 SUBTRACT X FROM 100 GIVING VAL
               DIVIDE 2 INTO VAL.
           IF VAL > P-HOLD GO TO PEN-4.
           IF X+P-HOLD > 100 GO TO REJ.
           IF VAL < P-HOLD AND F - P-HOLD > 0 GO TO PEN-4.
           IF VAL ≥ P-HOLD + 2 GO TO PEN-4 .
       REJ.
           MOVE P-HOLD TO VAL GO TO M3.
       PEN-4.
           SUBTRACT 1 FROM D.
           GO TO M3.
       PEN-3.
           IF VAL > X/2 COMPUTE VAL ROUNDED = X/2.
           SUBTRACT VAL FROM 0 GIVING VAL.
           IF VAL > P-HOLD GO TO PEN-6.
           IF D=4 AND F - P-HOLD > 0 GO TO PEN-6.
           IF X > 90 GO TO PEN-5.
           IF X + P-HOLD > 100 GO TO PEN-5.
           IF F - P-HOLD < 0 GO TO PEN-5.
           IF P-HOLD - VAL < 7 GO TO PEN-6.
       PEN-5.
           SUBTRACT 1 FROM D GO TO M3.
       PEN-6.
           MOVE P-HOLD TO VAL.
           GO TO M3.
       RNG.
           MOVE RN TO RN-MULT.
           MULTIPLY RN-MULT BY K GIVING RANDOM-NUMBER.
       PIN.
           PERFORM RNG.
           IF RND > 49 MOVE PHI-INV(RND - 49) TO PINV
           ELSE SUBTRACT PHI-INV(50 - RND) FROM 0 GIVING PINV.
       TO-PR.
       FGP.
           IF X<55 MOVE 000 TO PR.
           IF X≤65 MOVE 050 TO PR
           ELSE COMPUTE PR = 925 - (100 - X)*25.
       FG.
           PERFORM FGP.
           PERFORM RNG.
           IF PR > RNA ADD 3 TO SCO(T) MOVE 1 TO SW1.
       COM.
           PERFORM PIN.
           MULTIPLY SIG BY PINV GIVING SIGPINV.
           COMPUTE VAL ROUNDED = SIGPINV + MU.
       CPOS.
           IF T=1 MOVE 2 TO T ELSE MOVE 1 TO T.
           SUBTRACT X FROM 100 GIVING X.
           MOVE 1 TO D.
           MOVE 10 TO F.
```

Figure 10. Listing of the COBOL Simulation Model (Continued)

```
K-OFF.
    MOVE 28.7 TO MU.
    IF T=1 MOVE 17.6 TO SIG ELSE MOVE 7.9 TO SIG.
    PERFORM COM.
    IF VAL>0 MOVE VAL TO X ELSE MOVE 20 TO X.
    MOVE 10 TO F.
    MOVE 16 TO TIM PERFORM TIMER.
COIN-TOSS.
    PERFORM RNG.
    IF RND > 49 MOVE 1 TO T ELSE MOVE 2 TO T.
    IF T=1 MOVE 2 TO SEC-HALF ELSE MOVE 1 TO SEC-HALF.
ADV.
    MOVE SPACES TO VEC. WRITE VEC.
HDR1.
    MOVE "                  TEAM           YDG      LOSS
-   "              CHNG " TO H-VEC-B.
    MOVE "                 TIME WITH     BALL FOR       OR   PENAL
-   "TY  SPCL  SCORE   OF  " TO H-VEC-B.
    MOVE "                 LEFT BALL DOWN  ON  1ST DCSN GAIN AG YDG
-   " AC COND  1   2  POSS " TO H-VEC-B.
TIMER.
    IF SEC - TIM<0 AND MIN - 1 <0 MOVE "EOQ" TO EOQ.
    IF SEC - TIM < 0 SUBTRACT 1 FROM MIN ADD 60 TO
         SEC.
    SUBTRACT TIM FROM SEC.
Q-END.
    IF EOQ="EOQ" MOVE 15 TO MIN MOVE 00 TO SEC.
    MOVE SPACES TO EOQ.
    ADD 1 TO Q.
    IF Q=2 OR 4 GO TO L8.
    IF Q=3 MOVE 1 TO D MOVE SEC-HALF TO T PERFORM K-OFF GO TO L8.
    MOVE SCO(1) TO S-TEAM(SIMU,REP,1).
    ADD SCO(1) TO S-AVG(SIMU,1).
    MOVE SCO(2) TO S-TEAM(SIMU,REP,2).
    ADD SCO(2) TO S-AVG(SIMU,2).
    GO TO L4.
GAINER.
    IF VAL < 0 MOVE "-" TO LOSS.
    MOVE VAL TO ABS-D.
1ST-D.
    MOVE 10 TO F MOVE 1 TO D.
EOQ-TEST.
    IF EOQ = "EOQ" GO TO Q-END.
    GO TO L8.
PSEUDO-READ.
    ADD 1 TO H1. IF H1>H GO TO NULL.
    MOVE PA(H1) TO PARAMETERS-IP.
SUMMARIZATION.
    MOVE SPACES TO VEC.
    WRITE VEC BEFORE ADVANCING TO CHANNEL 1.
    MOVE G-TEAM(DIMU) TO GP.
    WRITE VEC. PERFORM ADV 2 TIMES.
    MOVE " *XA"      TO GRAPH-SYM  PERFORM GRAPH-FILL THRU GR2.
    MOVE "X ACTUAL SCORE  " TO ST1 WRITE VEC.
    MOVE "A SIMULATION AVG" TO ST1 WRITE VEC.
    MOVE "* SIMULATED SCORE   " TO ST1 WRITE VEC.
    PERFORM ADV.
    MOVE "  X  A   *" TO ST2 WRITE VEC.
    PERFORM ST-1 THRU ST-3 VARYING T FROM 1 BY 1 UNTIL T=2.
ST-1.
    IF T=1 MOVE "      GEORGIA TECH " TO ST1
      ELSE MOVE "         OPPONENTS " TO ST1   .
```

Figure 10. Listing of the COBOL Simulation Model (Continued)

```
           MOVE G-SCORE(DIMU,T) TO ST3(1).
           COMPUTE XA ROUNDED = S-AVG(SIMU,T)/N.
           MOVE XA TO ST3(2).
           PERFORM ST-2 VARYING XA FROM 1 BY 1 UNTIL XA=25.
       ST-2.
           MOVE S-TEAM(SIMU,XA,T) TO ST3(XA+2).
       ST-3.
           WRITE VEC MOVE SPACES TO VEC.
       SUM-2.
           MOVE SPACES TO VEC.
           PERFORM RANKING THRU S-R-2.
           PERFORM BIVARIATE-NORMAL THRU BN2.
           MOVE ZEROES TO SIMULATED-SCORES.
       HOR2.
           MOVE SPACES TO VEC.
           MOVE G-TEAM(SIMU) TO GP.
           MOVE "          SIMULATION" TO OPPOSITION.
           MOVE REP  TO FIN-SCO(1).
           WRITE VEC BEFORE ADVANCING 2 LINES.
           MOVE SPACES TO VEC.
           MOVE "             QUARTER" TO OPPOSITION.
           MOVE   Q  TO FIN-SCO(1).
           WRITE VEC BEFORE ADVANCING 3 LINES.
       SCOREKEEPER.
           MOVE SCO(1) TO SCO-O(1) MOVE SCO(2) TO SCO-O(2).
       GRAPH-FILL.
           PERFORM G-F-1 VARYING T FROM 1 BY 1 UNTIL T=N.
       G-F-1.
           COMPUTE XA ROUNDED =S-TEAM(SIMU,T,2)/2 + 1.
           COMPUTE XB ROUNDED =S-TEAM(SIMU,T,1)/2 + 1.
           IF XA>26 MOVE 26 TO XA.
           IF XB>26 MOVE 26 TO XB.
           IF X-CORD(XA,XB)≠ "  " MOVE "**" TO X-CORD(XA,XB) ELSE
           MOVE BORK TO X-CORD(XA,XB).
       G-F-2.
           COMPUTE XA ROUNDED=G-SCORE(DIMU,2)/2 + 1.
           COMPUTE XB ROUNDED=G-SCORE(DIMU,1)/2 + 1.
           MOVE CORK TO X-CORD(XA,XB).
       G-F-3.
           COMPUTE XA ROUNDED = S-AVG(SIMU,2)/(2*N) + 1.
           COMPUTE XB ROUNDED = S-AVG(SIMU,1)/(2*N) + 1.
           MOVE EORK TO X-CORD(XA,XB).
           PERFORM GR1 VARYING T FROM 1 BY 1 UNTIL T=26.
       GR1.
           MOVE Y-CORD(27 - T) TO GR-ROW.
           IF T=13 MOVE "POINTS SCORED" TO GP.
           IF T=14 MOVE " BY OPPONENTS" TO GP.
           IF T=1 OR 6 OR 11 OR 16 OR 21 OR 26 COMPUTE SKO=52 - 2*T.
               WRITE VEC MOVE SPACES TO VEC.
       GR2.
           MOVE " 0        10        20        30        40        50"
           TO GR-ROW. WRITE VEC.
           PERFORM ADV.
           MOVE SPACES TO GRAPH.
           MOVE "   POINTS SCORED BY GEORGIA TECH" TO GP2. WRITE VEC.
               PERFORM ADV 3 TIMES.
       RANKING.
           MOVE G-SCORE(DIMU,1) TO S-TEAM(SIMU,N+1,1).
           MOVE G-SCORE(DIMU,2) TO S-TEAM(SIMU,N+1,2).
           PERFORM R1 THRU R6 VARYING T FROM 1 BY 1 UNTIL T=2.
       R1.
           MOVE ZERO TO RA.
```

Figure 10. Listing of the COBOL Simulation Model (Continued)

```
           MOVE 1 TO R.
           PERFORM R2 THRU R6 VARYING XA FROM 0 BY 1 UNTIL R>N+1.
       R2.
           PERFORM R3 VARYING XB FROM 1 BY 1 UNTIL XB=N+1.
       R3.
           IF S-TEAM(SIMU,XB,T)=XA MOVE R TO S-RANK(SIMU,XB,T)
               ADD RA TO FACT ADD 1 TO RA ADD R TO FACT.
       R4.
           IF RA>1 DIVIDE RA INTO FACT ELSE GO TO R6.
           COMPUTE TIE(T)=TIE(T) - (RA**3 - RA)/12.
           PERFORM R5 VARYING XB FROM 1 BY 1 UNTIL XB=N+1.
       R5.
           IF S-TEAM(SIMU,XB,T)=XA MOVE FACT TO S-RANK(SIMU,XB,T).
       R6.
           ADD RA TO R.  MOVE ZERO TO RA.  MOVE ZEROES TO FACT.
       PRINT-RANK.
           WRITE VEC BEFORE ADVANCING TO CHANNEL 1.
       SPEARMANS-RHO.
           COMPUTE XSQ=(N**3 - N)/12   +TIE(1). MOVE ZEROES TO TIE(1).
           COMPUTE YSQ=(N**3 - N)/12   +TIE(2). MOVE ZEROES TO TIE(2).
           MOVE ZEROES TO DSQ.
           PERFORM S-R-1 VARYING XB FROM 1 BY 1 UNTIL XB=N+1.
       S-R-1.
           COMPUTE DSQ=DSQ+(S-RANK(SIMU,XB,1) - S-RANK(SIMU,XB,2))**2.
       S-R-2.
           MOVE "SPEARMANS RHO " TO GP.
           COMPUTE RHO=(XSQ + YSQ - DSQ)/(2*SQRT(XSQ*YSQ)).
           MOVE RHO TO GP3.
           WRITE VEC MOVE SPACES TO VEC.
           MOVE S-RANK(SIMU,N+1,1) TO BRACER WRITE VEC.
           MOVE S-RANK(SIMU,N+1,2) TO BRACER WRITE VEC.
       BIVARIATE-NORMAL.
           MOVE ZEROES TO XY.
           MOVE ZEROES TO XSQ.
           MOVE ZEROES TO YSQ.
           PERFORM BN1 VARYING XA FROM 1 BY 1 UNTIL XA=N.
       BN1.
           COMPUTE XSQ=XSQ+(S-TEAM(SIMU,XA,1))**2.
           COMPUTE YSQ=YSQ+(S-TEAM(SIMU,XA,2))**2.
           COMPUTE XY=XY+S-TEAM(SIMU,XA,1)*S-TEAM(SIMU,XA,2).
       BN2.
           MOVE XSQ TO BRACER WRITE VEC.
           MOVE YSQ TO BRACER WRITE VEC.
           MOVE XY  TO BRACER WRITE VEC.
           DIVIDE N INTO S-AVG(SIMU,1).
           DIVIDE N INTO S-AVG(SIMU,2).
           COMPUTE S-VAR(SIMU,1)=(XSQ - N*(S-AVG(SIMU,1)**2))/(N - 1).
           MOVE S-VAR(SIMU,1) TO BRACER WRITE VEC.
           COMPUTE S-VAR(SIMU,2)=(YSQ - N*(S-AVG(SIMU,2)**2))/(N - 1).
           MOVE S-VAR(SIMU,2) TO BRACER WRITE VEC.
           COMPUTE COVAR=XY/N - S-AVG(SIMU,1)*S-AVG(SIMU,2).
           MOVE COVAR TO BRACER WRITE VEC.
           COMPUTE CORR=COVAR/SQRT(S-VAR(SIMU,1)*S-VAR(SIMU,2)).
           MOVE CORR TO BRACER WRITE VEC.
           COMPUTE SIGMA=(((G-SCORE(DIMU,1) - S-AVG(SIMU,1))**2/S-VAR(SI
      -    MU,1)) - (2*CORR*(G-SCORE(DIMU,1) - S-AVG(SIMU,1))*(G-SCORE(D
      -    IMU,2) - S-AVG(SIMU,2)))/(SQRT(S-VAR(SIMU,1))*SQRT(S-VAR(SIM
      -    U,2)))+((G-SCORE(DIMU,2) - S-AVG(SIMU,2))**2/S-VAR(SIMU,2)))/
      -    (1 - CORR**2).
           MOVE SIGMA TO BRACER WRITE VEC.
           COMPUTE SIGMA= SQRT(SIGMA).
           MOVE SIGMA TO BRACER WRITE VEC.
```

Figure 10. Listing of the COBOL Simulation Model (Continued)

```
           COMPUTE S-VAR(SIMU,1)=SQRT(S-VAR(SIMU,1)).
           COMPUTE S-VAR(SIMU,2)=SQRT(S-VAR(SIMU,2)).
           MOVE S-VAR(SIMU,1) TO BRACER WRITE VEC.
           MOVE S-VAR(SIMU,2) TO BRACER WRITE VEC.
           WRITE VEC BEFORE ADVANCING TO CHANNEL 1.
       C-POS.
           IF T=1 MOVE 2 TO T ELSE MOVE 1 TO T.
           SUBTRACT X FROM 100 GIVING X.
           MOVE 1 TO D.
           MOVE 10 TO F.
       NULL.
           CLOSE OP-FILE. STOP RUN.
       END-OF-JOB.
```

Figure 10. Listing of the COBOL Simulation Model (Continued)

GAME PARAMETERS

GEORGIA TECH

PR(INCOMPLETE PASS)	.310
PR(PASS NOT THROWN)	.000

	MEAN	STANDARD DEVIATION
RUN	+4.94	7.97
PASS POS	+12.00	2.87
PASS NEG	+.00	.00

FLORIDA 1960

PR(INCOMPLETE PASS)	.390
PR(PASS NOT THROWN)	.180

	MEAN	STANDARD DEVIATION
RUN	+4.41	4.52
PASS POS	+11.91	8.22
PASS NEG	-8.75	3.94

DECISION PARAMETERS

GEORGIA TECH

		PERCENT OF PLAYS THAT ARE:			
		RUN	PASS	PUNT	FG
DOWN	1	87	13	00	00
DOWN	2	61	39	00	00
DOWN	3	13	37	37	13
DOWN	4	00	00	1 00	00

FLORIDA 1960

		PERCENT OF PLAYS THAT ARE:			
		RUN	PASS	PUNT	FG
DOWN	1	64	36	00	00
DOWN	2	70	26	04	00
DOWN	3	46	46	08	00
DOWN	4	29	00	29	42

Figure 11. Input to the Simulation Model for the Georgia Tech-Florida Game, 1960

GAME PARAMETERS

GEORGIA TECH

PR(INCOMPLETE PASS) .416

PR(PASS NOT THROWN) .143

	MEAN	STANDARD DEVIATION
RUN	+3.50	6.16
PASS POS	+11.57	5.09
PASS NEG	-2.00	.00

ALABAMA 1960

PR(INCOMPLETE PASS) .400

PR(PASS NOT THROWN) .130

	MEAN	STANDARD DEVIATION
RUN	+3.48	3.80
PASS POS	+14.75	8.22
PASS NEG	-6.67	3.21

DECISION PARAMETERS

GEORGIA TECH

		PERCENT OF PLAYS THAT ARE:			
		RUN	PASS	PUNT	FG
DOWN	1	78	22	00	00
DOWN	2	56	44	00	00
DOWN	3	67	13	20	00
DOWN	4	14	00	72	14

ALABAMA 1960

		PERCENT OF PLAYS THAT ARE:			
		RUN	PASS	PUNT	FG
DOWN	1	61	35	00	04
DOWN	2	59	41	00	00
DOWN	3	18	55	27	00
DOWN	4	44	28	28	00

Figure 12. Input to the Simulation Model for the Georgia Tech-Alabama Game, 1960

GAME PARAMETERS

GEORGIA TECH

PR(INCOMPLETE PASS)	.580
PR(PASS NOT THROWN)	.670

	MEAN	STANDARD DEVIATION
RUN	+3.23	3.36
PASS POS	+12.17	6.30
PASS NEG	-2.00	.00

TULANE 1960

PR(INCOMPLETE PASS)	.390
PR(PASS NOT THROWN)	.140

	MEAN	STANDARD DEVIATION
RUN	+2.88	2.92
PASS POS	+14.75	9.88
PASS NEG	-8.00	4.58

DECISION PARAMETERS

GEORGIA TECH

		PERCENT OF PLAYS THAT ARE:			
		RUN	PASS	PUNT	FG
DOWN	1	79	17	04	00
DOWN	2	65	35	00	00
DOWN	3	50	29	21	00
DOWN	4	00	00	80	20

TULANE 1960

		PERCENT OF PLAYS THAT ARE:			
		RUN	PASS	PUNT	FG
DOWN	1	66	34	00	00
DOWN	2	76	24	00	00
DOWN	3	23	62	15	00
DOWN	4	12	13	75	00

Figure 13. Input to the Simulation Model for the Georgia Tech-Tulane Game, 1960

GAME PARAMETERS

GEORGIA TECH

PR(INCOMPLETE PASS)	.250
PR(PASS NOT THROWN)	.080

	MEAN	STANDARD DEVIATION
RUN	+4.23	4.20
PASS POS	+14.00	9.96
PASS NEG	-1.00	.00

STHRN CAL 1961

PR(INCOMPLETE PASS)	.400
PR(PASS NOT THROWN)	.080

	MEAN	STANDARD DEVIATION
RUN	+1.84	3.91
PASS POS	+15.87	12.81
PASS NEG	-3.00	1.41

DECISION PARAMETERS

GEORGIA TECH

		PERCENT OF PLAYS THAT ARE:			
		RUN	PASS	PUNT	FG
DOWN	1	80	20	00	00
DOWN	2	86	14	00	00
DOWN	3	53	39	08	00
DOWN	4	00	00	33	67

STHRN CAL 1961

		PERCENT OF PLAYS THAT ARE:			
		RUN	PASS	PUNT	FG
DOWN	1	50	50	00	00
DOWN	2	67	33	00	00
DOWN	3	31	69	00	00
DOWN	4	22	33	45	00

Figure 14. Input to the Simulation Model for the Georgia Tech-Southern California Game, 1961

GAME PARAMETERS

GEORGIA TECH

PR(INCOMPLETE PASS) .610

PR(PASS NOT THROWN) .280

	MEAN	STANDARD DEVIATION
RUN	+3.51	3.49
PASS POS	+24.80	18.17
PASS NEG	-7.40	5.20

TULANE 1961

PR(INCOMPLETE PASS) .730

PR(PASS NOT THROWN) .170

	MEAN	STANDARD DEVIATION
RUN	+3.77	5.86
PASS POS	+10.20	2.16
PASS NEG	-6.50	4.20

DECISION PARAMETERS

GEORGIA TECH

PERCENT OF PLAYS THAT ARE:

		RUN	PASS	PUNT	FG
DOWN	1	80	10	10	00
DOWN	2	70	25	05	00
DOWN	3	39	55	06	00
DOWN	4	00	25	75	00

TULANE 1961

PERCENT OF PLAYS THAT ARE:

		RUN	PASS	PUNT	FG
DOWN	1	59	36	05	00
DOWN	2	56	44	00	00
DOWN	3	31	54	15	00
DOWN	4	12	13	75	00

Figure 15. Input to the Simulation Model for the Georgia Tech-Tulane Game, 1961

GAME PARAMETERS

GEORGIA TECH

PR(INCOMPLETE PASS) .280

PR(PASS NOT THROWN) .220

	MEAN	STANDARD DEVIATION
RUN	+4.43	5.76
PASS POS	+12.80	6.04
PASS NEG	-7.33	2.65

RICE 1961

PR(INCOMPLETE PASS) .660

PR(PASS NOT THROWN) .000

	MEAN	STANDARD DEVIATION
RUN	+4.59	5.22
PASS POS	+10.80	6.97
PASS NEG	+.00	.00

DECISION PARAMETERS

GEORGIA TECH

		PERCENT OF PLAYS THAT ARE:			
		RUN	PASS	PUNT	FG
DOWN	1	61	39	00	00
DOWN	2	60	40	00	00
DOWN	3	65	35	00	00
DOWN	4	20	00	60	20

RICE 1961

		PERCENT OF PLAYS THAT ARE:			
		RUN	PASS	PUNT	FG
DOWN	1	75	25	00	00
DOWN	2	63	37	00	00
DOWN	3	33	56	11	00
DOWN	4	25	00	75	00

Figure 16. Input to the Simulation Model for the Georgia Tech-Rice Game, 1961

GAME PARAMETERS

GEORGIA TECH

PR(INCOMPLETE PASS) .500

PR(PASS NOT THROWN) .000

	MEAN	STANDARD DEVIATION
RUN	+7.19	6.48
PASS POS	+18.83	6.11
PASS NEG	+.00	.00

TULANE 1962

PR(INCOMPLETE PASS) .570

PR(PASS NOT THROWN) .130

	MEAN	STANDARD DEVIATION
RUN	+3.63	3.31
PASS POS	+10.67	2.73
PASS NEG	-8.50	7.77

DECISION PARAMETERS

GEORGIA TECH

		PERCENT OF PLAYS THAT ARE:			
		RUN	PASS	PUNT	FG
DOWN	1	70	30	00	00
DOWN	2	89	11	00	00
DOWN	3	17	67	16	00
DOWN	4	33	00	33	34

TULANE 1962

		PERCENT OF PLAYS THAT ARE:			
		RUN	PASS	PUNT	FG
DOWN	1	76	24	00	00
DOWN	2	67	33	00	00
DOWN	3	59	33	08	00
DOWN	4	14	00	86	00

Figure 17. Input to the Simulation Model for the Georgia Tech-Tulane Game, 1962

GAME PARAMETERS

GEORGIA TECH

PR(INCOMPLETE PASS) .480

PR(PASS NOT THROWN) .220

	MEAN	STANDARD DEVIATION
RUN	+3.44	3.84
PASS POS	+12.64	7.07
PASS NEG	-4.83	2.99

AUBURN 1962

PR(INCOMPLETE PASS) .430

PR(PASS NOT THROWN) .000

	MEAN	STANDARD DEVIATION
RUN	+3.98	8.63
PASS POS	+12.30	6.95
PASS NEG	-8.50	7.77

DECISION PARAMETERS

GEORGIA TECH

PERCENT OF PLAYS THAT ARE:

		RUN	PASS	PUNT	FG
DOWN	1	79	21	00	00
DOWN	2	50	50	00	00
DOWN	3	07	80	13	00
DOWN	4	00	33	67	00

AUBURN 1962

PERCENT OF PLAYS THAT ARE:

		RUN	PASS	PUNT	FG
DOWN	1	83	17	00	00
DOWN	2	48	52	00	00
DOWN	3	35	36	29	00
DOWN	4	29	00	29	42

Figure 18. Input to the Simulation Model for the Georgia Tech-Auburn Game, 1962

GAME PARAMETERS

GEORGIA TECH

PR(INCOMPLETE PASS) .560

PR(PASS NOT THROWN) .000

	MEAN	STANDARD DEVIATION
RUN	+2.68	4.23
PASS POS	+10.25	7.63
PASS NEG	+.00	.00

FLORIDA 1963

PR(INCOMPLETE PASS) .500

PR(PASS NOT THROWN) .500

	MEAN	STANDARD DEVIATION
RUN	+2.97	4.08
PASS POS	+9.25	2.98
PASS NEG	-9.75	3.61

DECISION PARAMETERS

GEORGIA TECH

PERCENT OF PLAYS THAT ARE:

		RUN	PASS	PUNT	FG
DOWN	1	80	20	00	00
DOWN	2	88	12	00	00
DOWN	3	25	59	16	00
DOWN	4	33	00	11	56

FLORIDA 1963

PERCENT OF PLAYS THAT ARE:

		RUN	PASS	PUNT	FG
DOWN	1	80	20	00	00
DOWN	2	63	25	12	00
DOWN	3	33	63	04	00
DOWN	4	00	11	89	00

Figure 19. Input to the Simulation Model for the Georgia Tech-Florida Game, 1963

GAME PARAMETERS

GEORGIA TECH

PR(INCOMPLETE PASS) .390

PR(PASS NOT THROWN) .180

	MEAN	STANDARD DEVIATION
RUN	+4.53	6.02
PASS POS	+12.10	6.69
PASS NEG	-8.20	3.11

FLORIDA STATE 1963

PR(INCOMPLETE PASS) .750

PR(PASS NOT THROWN) .000

	MEAN	STANDARD DEVIATION
RUN	+4.89	4.96
PASS POS	+11.50	8.96
PASS NEG	+.00	.00

DECISION PARAMETERS

GEORGIA TECH

PERCENT OF PLAYS THAT ARE:

		RUN	PASS	PUNT	FG
DOWN	1	64	36	00	00
DOWN	2	39	61	00	00
DOWN	3	17	75	08	00
DOWN	4	11	00	67	22

FLORIDA STATE 1963

PERCENT OF PLAYS THAT ARE:

		RUN	PASS	PUNT	FG
DOWN	1	77	23	00	00
DOWN	2	71	29	00	00
DOWN	3	43	36	31	00
DOWN	4	00	15	85	00

Figure 20. Input to the Simulation Model for the Georgia Tech-Florida State Game, 1963

GAME PARAMETERS

GEORGIA TECH

PR(INCOMPLETE PASS)	.420
PR(PASS NOT THROWN)	.200

	MEAN	STANDARD DEVIATION
RUN	+2.15	4.08
PASS POS	+23.71	12.60
PASS NEG	-7.67	8.32

NAVY 1964

PR(INCOMPLETE PASS)	.650
PR(PASS NOT THROWN)	.210

	MEAN	STANDARD DEVIATION
RUN	+2.91	3.21
PASS POS	+10.67	7.34
PASS NEG	-5.57	3.30

DECISION PARAMETERS

GEORGIA TECH

PERCENT OF PLAYS THAT ARE:

		RUN	PASS	PUNT	FG
DOWN	1	76	24	00	00
DOWN	2	65	35	00	00
DOWN	3	50	40	10	00
DOWN	4	17	00	75	08

NAVY 1964

PERCENT OF PLAYS THAT ARE:

		RUN	PASS	PUNT	FG
DOWN	1	71	29	00	00
DOWN	2	47	48	05	00
DOWN	3	33	67	00	00
DOWN	4	00	43	50	07

Figure 21. Input to the Simulation Model for the Georgia Tech-Navy Game, 1964

GAME PARAMETERS

GEORGIA TECH

PR(INCOMPLETE PASS) .500

PR(PASS NOT THROWN) .200

	MEAN	STANDARD DEVIATION
RUN	+3.22	3.39
PASS POS	+16.50	6.36
PASS NEG	-1.00	.00

GEORGIA 1964

PR(INCOMPLETE PASS) .500

PR(PASS NOT THROWN) .333

	MEAN	STANDARD DEVIATION
RUN	+3.86	4.08
PASS POS	+19.00	15.71
PASS NEG	-9.67	2.51

DECISION PARAMETERS

GEORGIA TECH

PERCENT OF PLAYS THAT ARE:

		RUN	PASS	PUNT	FG
DOWN	1	93	07	00	00
DOWN	2	92	08	00	00
DOWN	3	78	22	00	00
DOWN	4	00	14	86	00

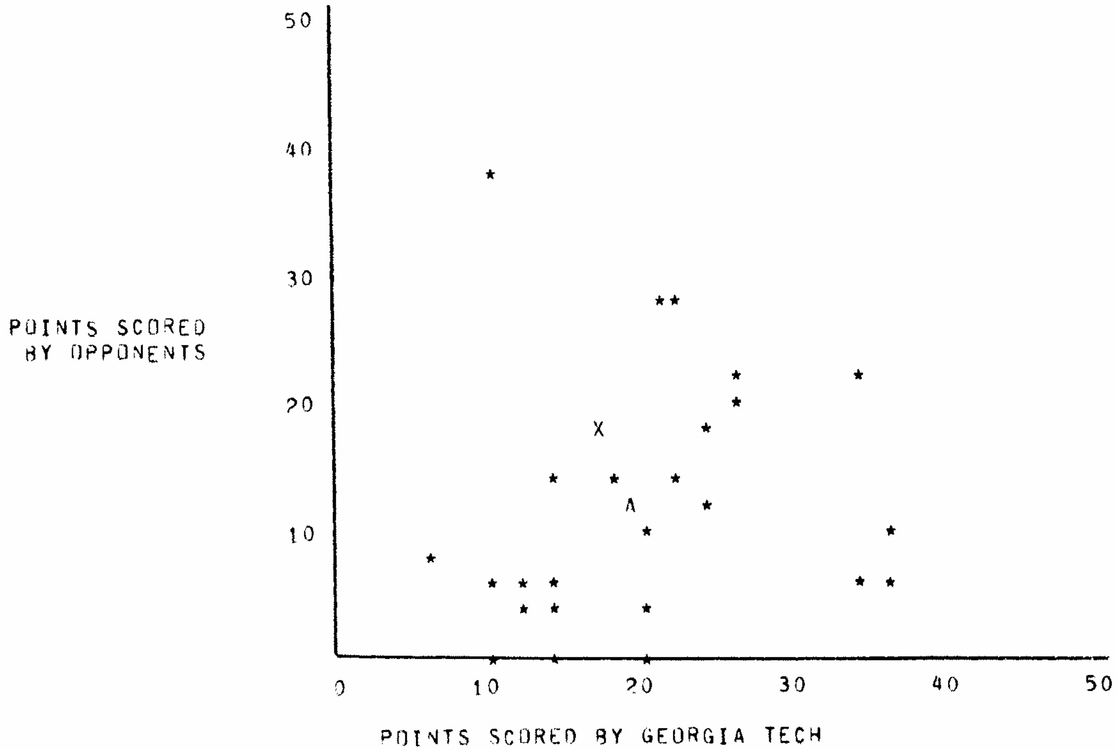
GEORGIA 1964

PERCENT OF PLAYS THAT ARE:

		RUN	PASS	PUNT	FG
DOWN	1	87	13	00	00
DOWN	2	71	29	00	00
DOWN	3	61	16	23	00
DOWN	4	00	00	00	1 00

Figure 22. Input to the Simulation Model for the Georgia Tech-Georgia Game, 1964

FLORIDA 1960

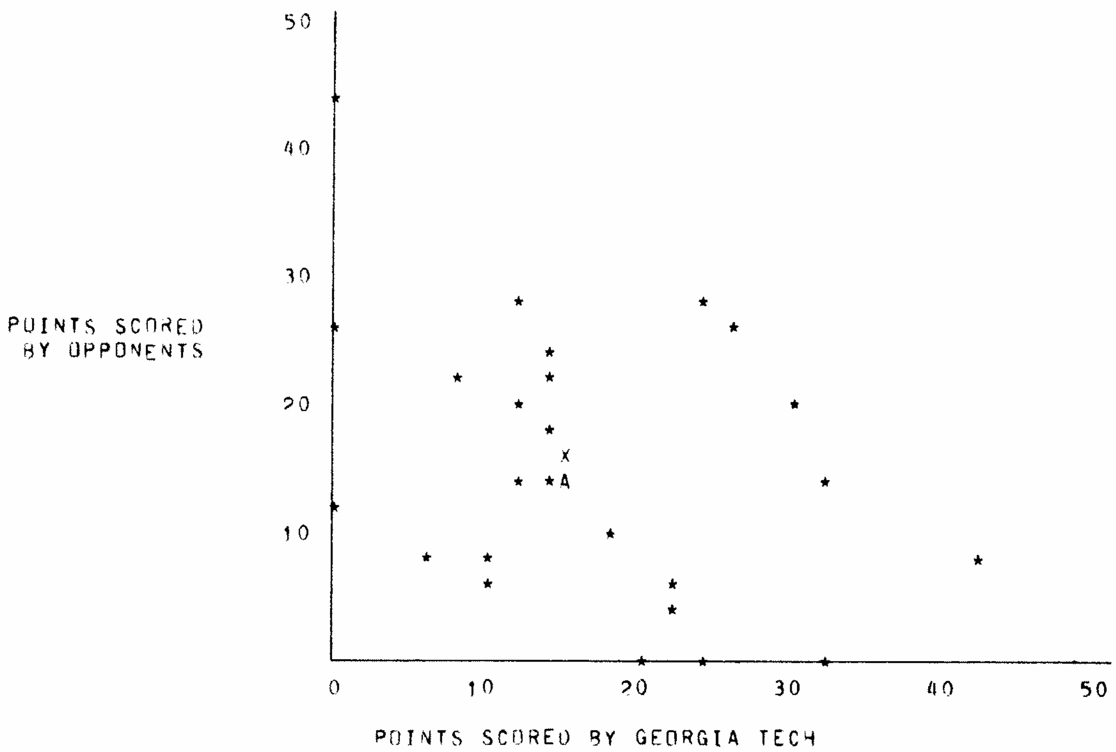


X ACTUAL SCORE
A SIMULATION AVG
* SIMULATED SCORE

	X	A	*																								
GEORGIA TECH	17	19	34	23	35	21	9	20	13	13	6	12	13	12	36	23	21	17	26	26	10	9	13	33	19	21	20
OPPONENTS	18	12	6	17	10	28	38	0	6	13	8	6	0	3	6	12	14	14	20	22	0	6	3	21	3	28	10

Figure 23. Output of the Simulation of the Georgia Tech-Florida Game, 1960

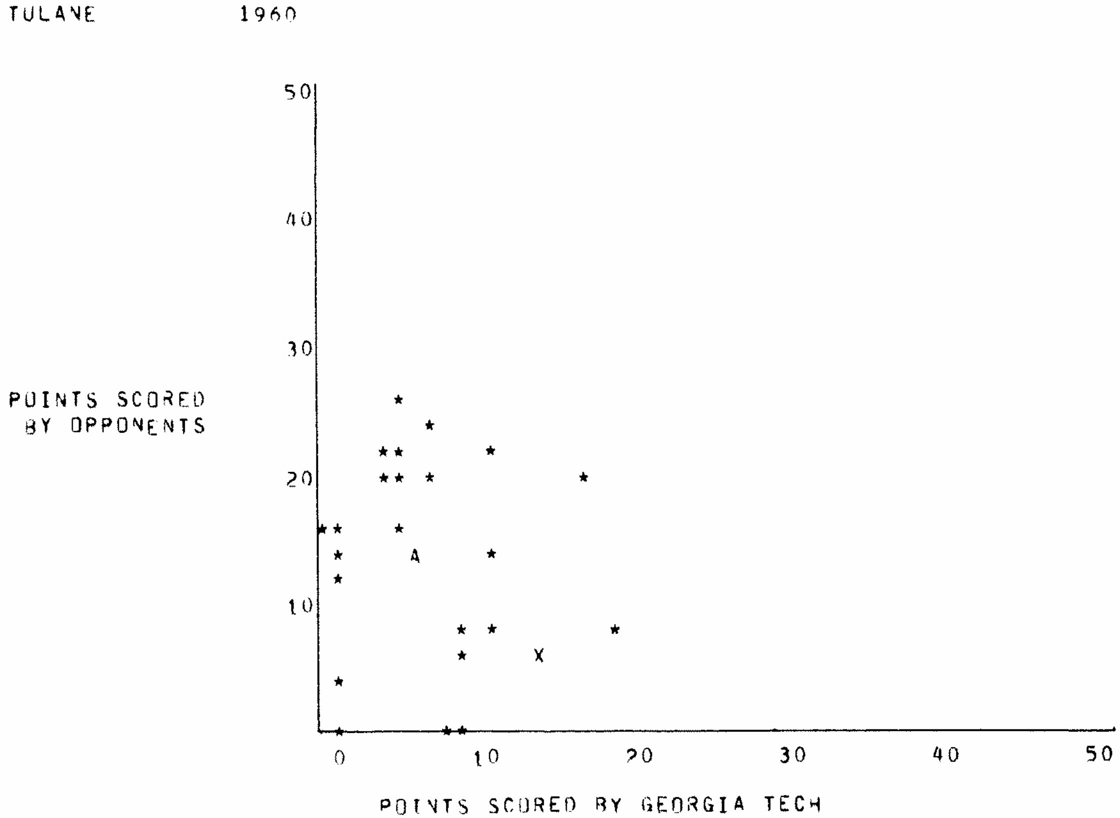
ALABAMA 1960



X ACTUAL SCORE
A SIMULATION AVG
* SIMULATED SCORE

	X	A	*																								
GEORGIA TECH	15	16	24	12	13	10	13	13	5	23	42	12	25	20	32	0	21	0	30	22	13	11	0	7	31	18	9
OPPONENTS	16	15	27	20	14	6	17	23	7	0	7	27	26	0	0	44	6	12	20	3	21	14	26	22	14	9	7

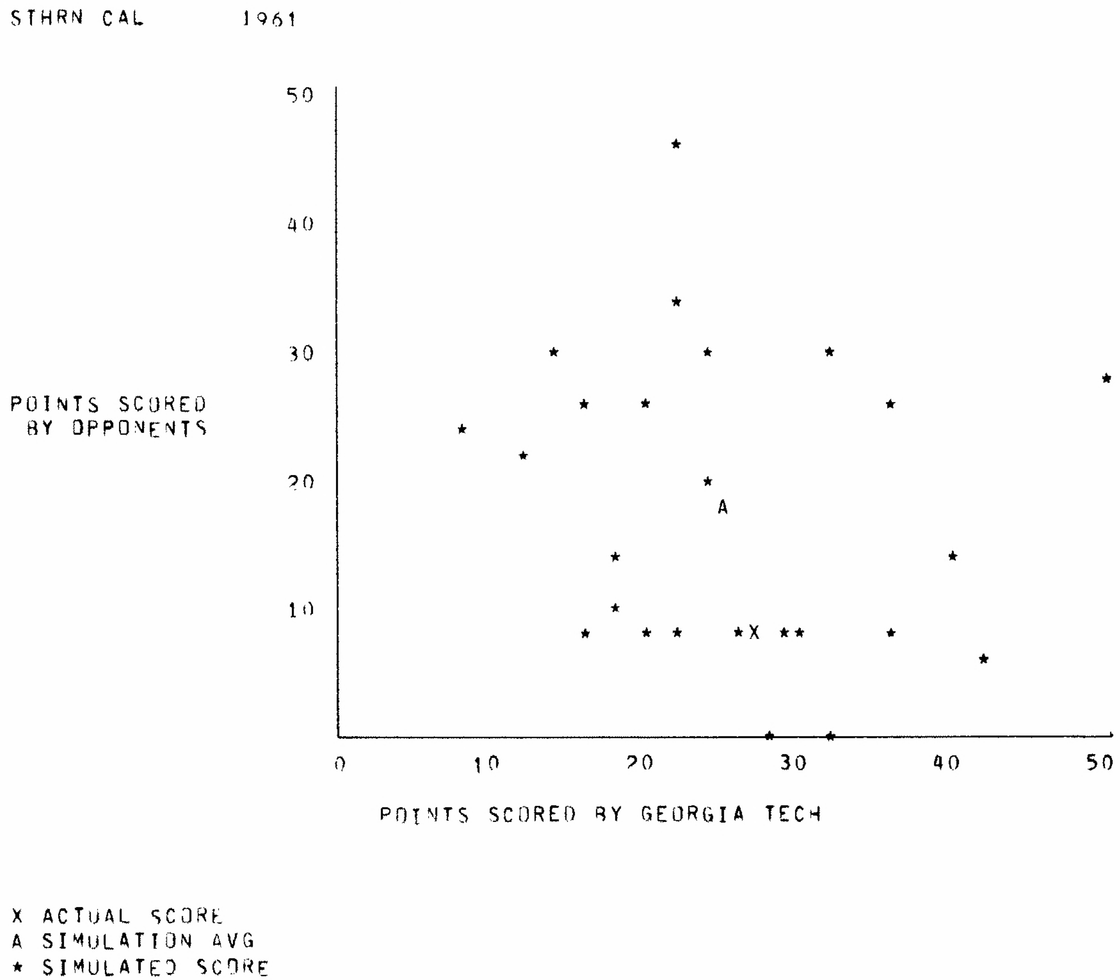
Figure 24. Output of the Simulation of the Georgia Tech-Alabama Game, 1960



X ACTUAL SCORE
A SIMULATION AVG
* SIMULATED SCORE

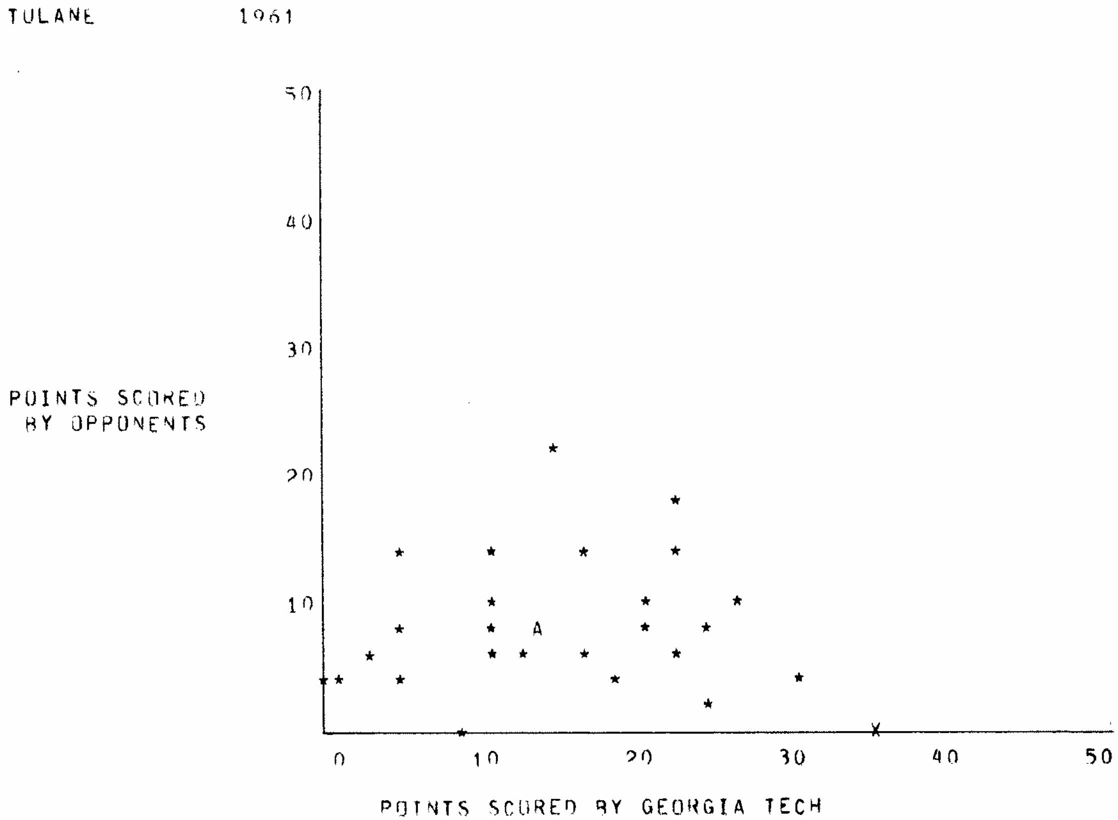
	X	A	*																								
GEORGIA TECH	14	5	8	10	0	0	17	3	0	0	3	3	15	3	10	6	0	8	0	3	6	6	7	8	3	3	9
OPPONENTS	6	14	0	13	16	3	8	15	0	16	22	21	20	20	22	20	13	7	12	25	13	23	6	0	20	21	7

Figure 25. Output of the Simulation of the Georgia Tech-Tulane Game, 1960



	X	A	*																								
GEORGIA TECH	27	25	30	39	26	32	22	18	35	23	13	16	49	27	36	12	31	21	19	19	30	16	23	17	42	22	7
OPPONENTS	7	17	7	14	7	29	46	14	7	20	29	7	27	0	26	21	0	33	7	26	8	26	29	9	6	7	24

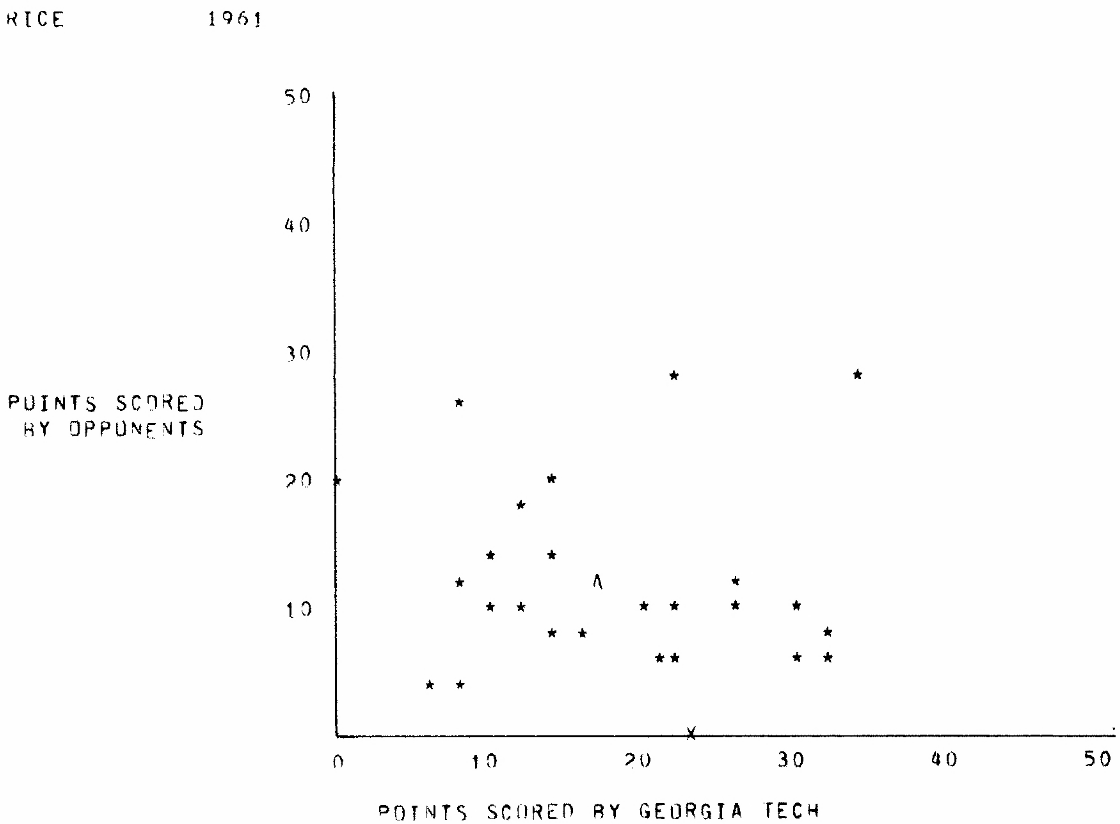
Figure 26. Output of the Simulation of the Georgia Tech-Southern California Game, 1961



X ACTUAL SCORE
A SIMULATION AVG
* SIMULATED SCORE

	X	A	*																								
GEORGIA TECH	35	14	22	3	22	24	3	0	20	29	3	2	10	16	12	18	26	13	9	9	20	7	10	16	21	23	0
OPPONENTS	0	8	6	14	14	7	4	3	7	3	7	6	6	6	6	3	9	22	13	7	9	0	10	14	17	2	3

Figure 27. Output of the Simulation of the Georgia Tech-Tulane Game, 1961

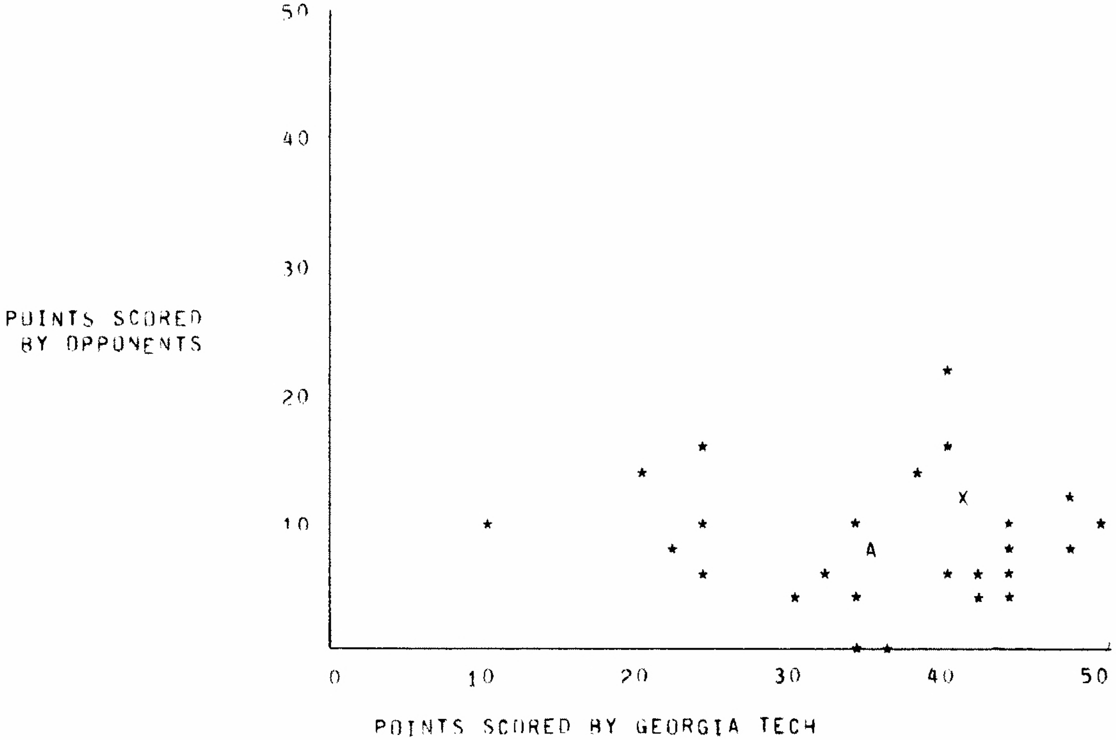


X ACTUAL SCORE
A SIMULATION AVG
* SIMULATED SCORE

	X	A	*																								
GEORGIA TECH	24	17	0	21	31	25	7	9	10	32	25	13	14	29	21	11	12	7	6	14	20	16	29	21	8	34	22
OPPONENTS	0	12	20	6	7	12	12	14	9	6	10	13	20	9	10	9	17	3	3	7	10	7	6	27	26	28	5

Figure 28. Output of the Simulation of the Georgia Tech-Rice Game, 1961

TULANE 1962

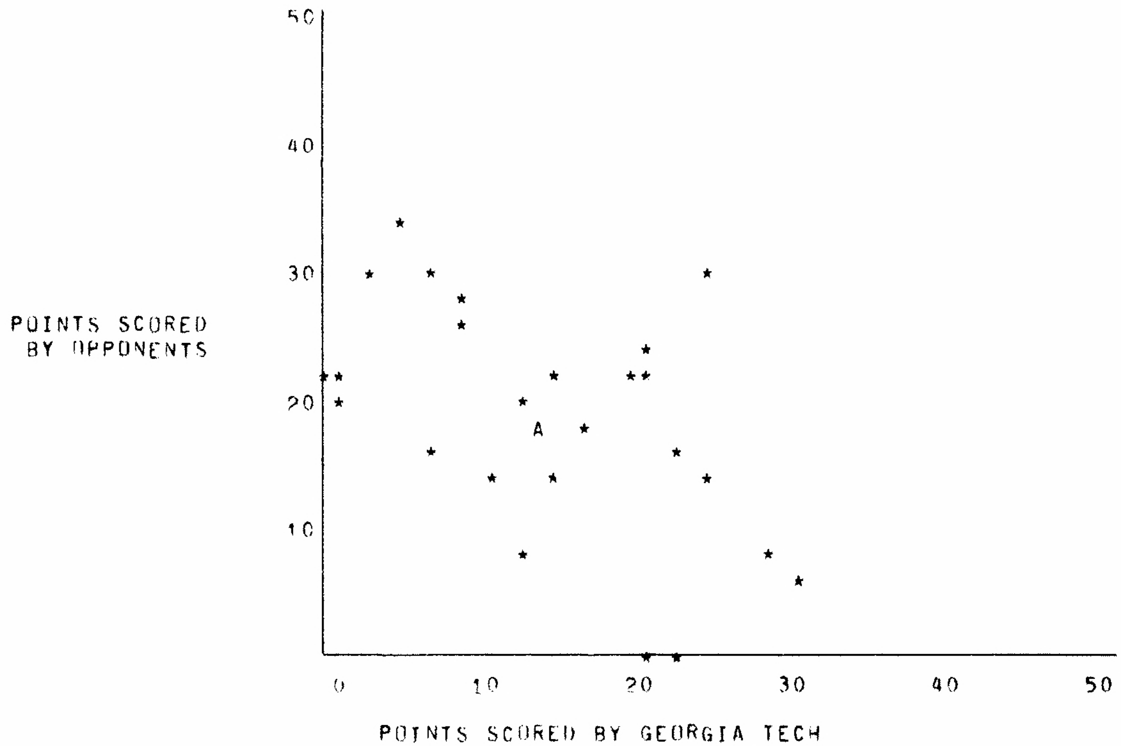


X ACTUAL SCORE
A SIMULATION AVG
* SIMULATED SCORE

	X	A	*																							
GEORGIA TECH	42	35	44	10	33	44	22	44	23	40	48	35	40	23	42	41	31	34	43	34	40	20	47	37	30	24	55
OPPONENTS	12	8	3	10	0	9	7	6	16	16	12	0	6	9	3	6	6	3	8	9	21	14	7	14	3	6	9

Figure 29. Output of the Simulation of the Georgia Tech-Tulane Game, 1962

AUBURN 1962

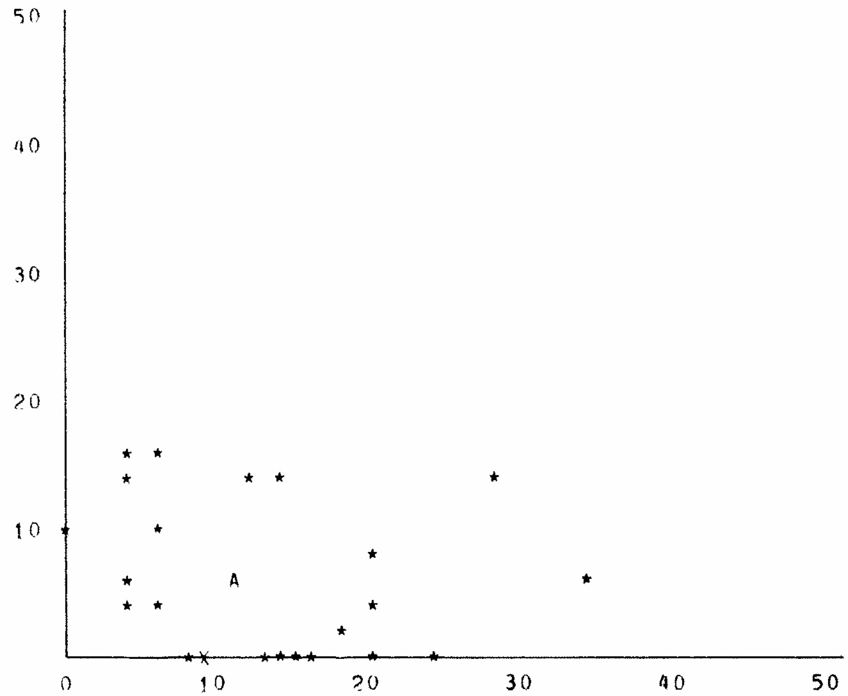


X ACTUAL SCORE
A SIMULATION AVG
* SIMULATED SCORE

	X	A	*																								
GEORGIA TECH	14	13	2	12	9	30	20	22	0	23	20	27	8	3	0	23	19	20	12	13	0	14	6	5	7	15	22
OPPONENTS	17	18	30	7	14	6	0	0	21	29	24	7	27	33	21	13	21	21	19	13	19	21	15	29	25	17	16

Figure 30. Output of the Simulation of the Georgia Tech-Auburn Game, 1962

FLORIDA 1963



POINTS SCORED BY OPPONENTS

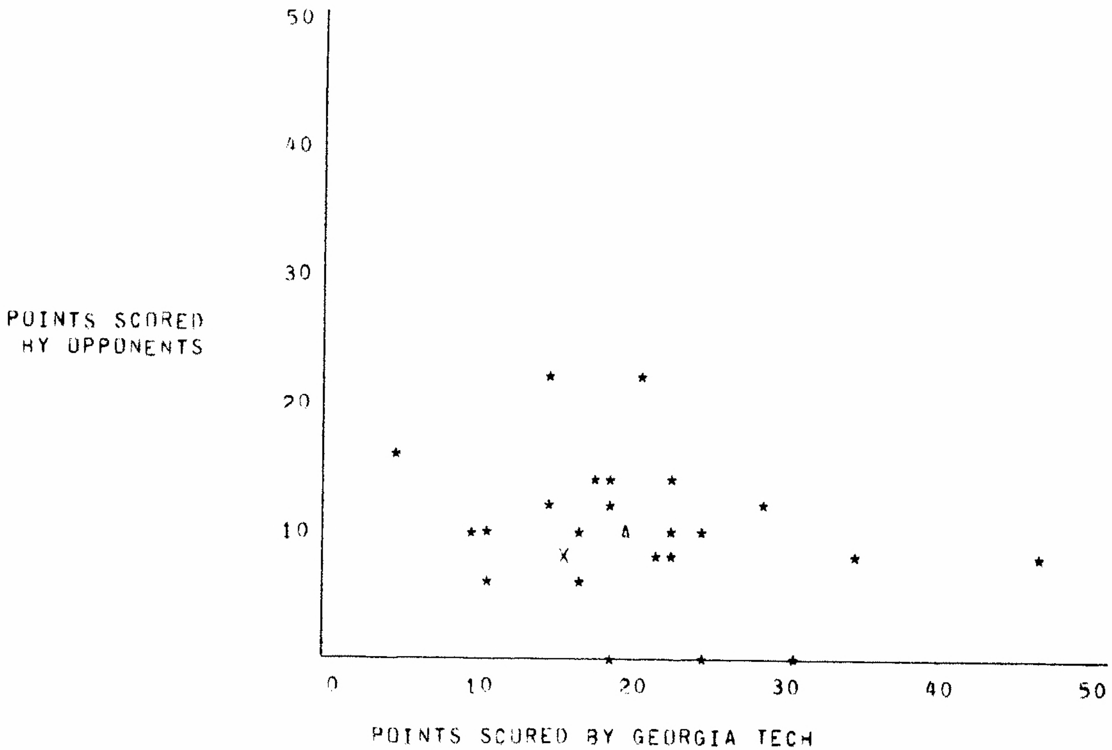
POINTS SCORED BY GEORGIA TECH

X ACTUAL SCORE
A SIMULATION AVG
* SIMULATED SCORE

	X	A	*																								
GEORGIA TECH	9	12	7	18	20	33	20	16	11	3	14	15	11	23	5	10	3	4	3	14	5	19	27	14	0	10	5
OPPONENTS	0	6	0	2	3	6	7	0	6	6	0	0	14	0	3	0	3	13	15	0	15	0	13	13	10	0	10

Figure 31. Output of the Simulation of the Georgia Tech-Florida Game, 1963

FLORIDA STATE 1963

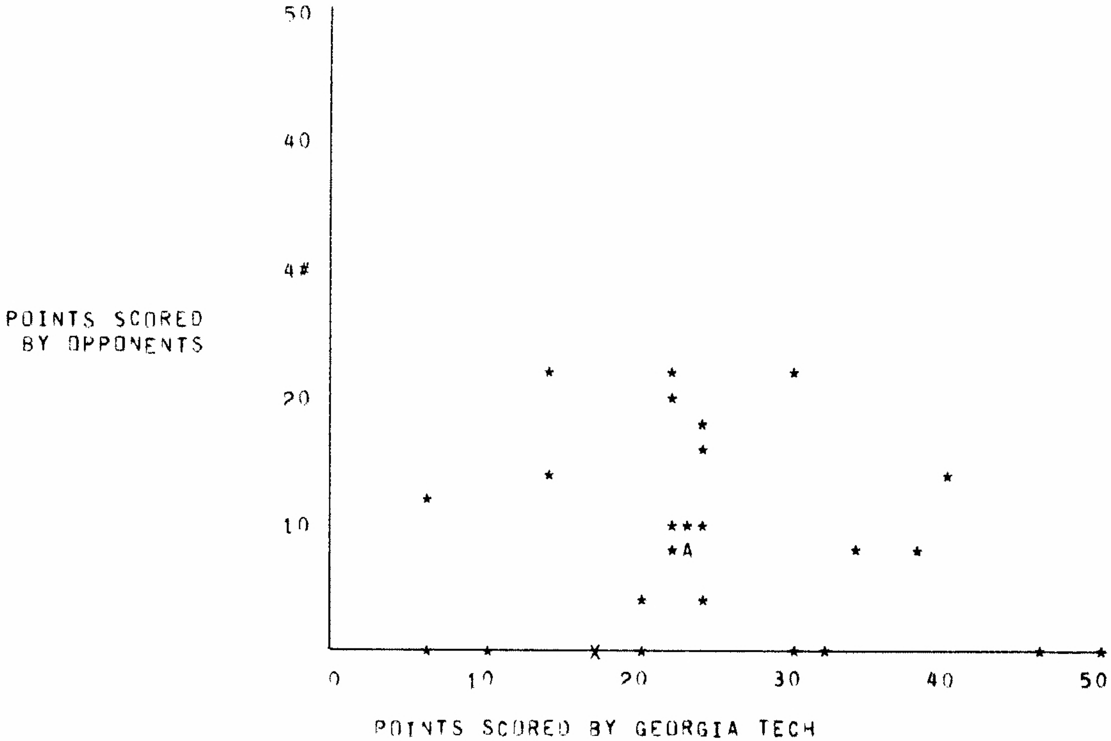


X ACTUAL SCORE
A SIMULATION AVG
* SIMULATED SCORE

	X	A	*																								
GEORGIA TECH	15	19	27	17	21	45	23	30	17	22	9	16	20	14	21	17	17	34	10	22	14	24	3	16	9	10	20
OPPONENTS	7	10	12	13	10	7	0	0	12	13	6	10	9	12	7	0	13	8	9	8	22	10	16	5	10	10	21

Figure 32. Output of the Simulation of the Georgia Tech-Florida State Game, 1963

NAVY 1964

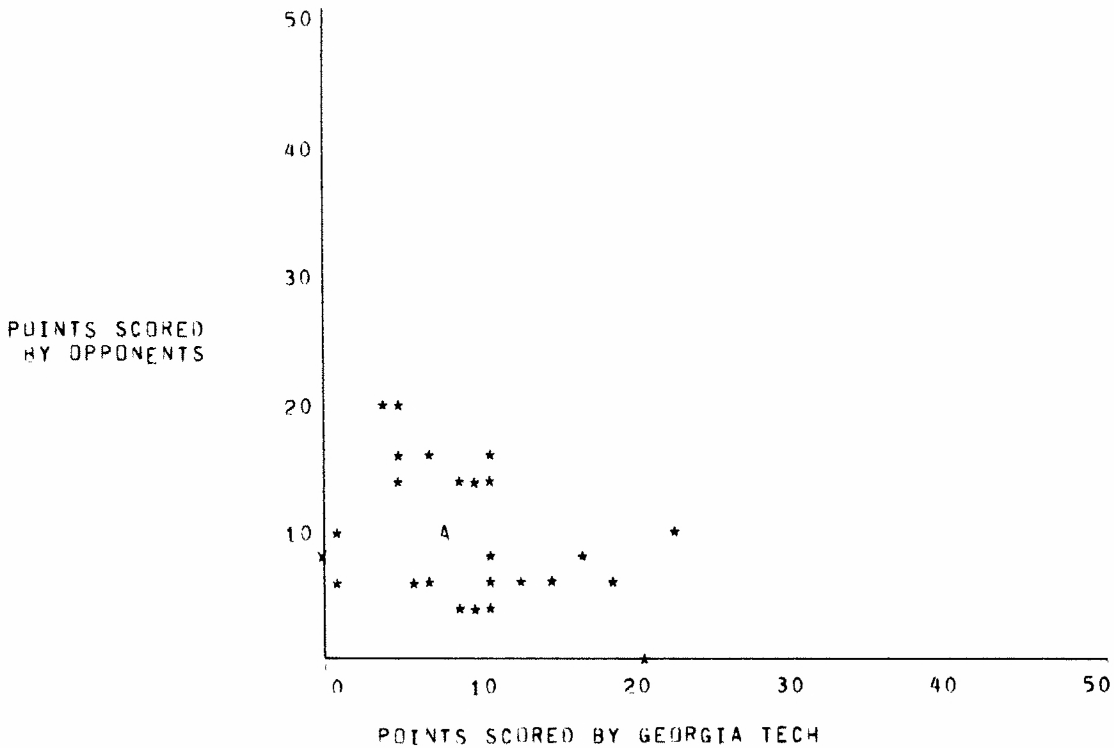


X ACTUAL SCORE
A SIMULATION AVG
* SIMULATED SCORE

	X	A	*																								
GEORGIA TECH	17	25	9	22	13	22	18	34	32	21	22	24	6	20	29	23	14	19	38	58	24	6	45	40	23	30	23
OPPONENTS	0	8	0	9	21	19	0	7	0	21	7	10	0	3	21	15	13	0	7	0	10	11	0	14	17	0	3

Figure 33. Output of the Simulation of the Georgia Tech-Navy Game, 1964

GEORGIA 1964

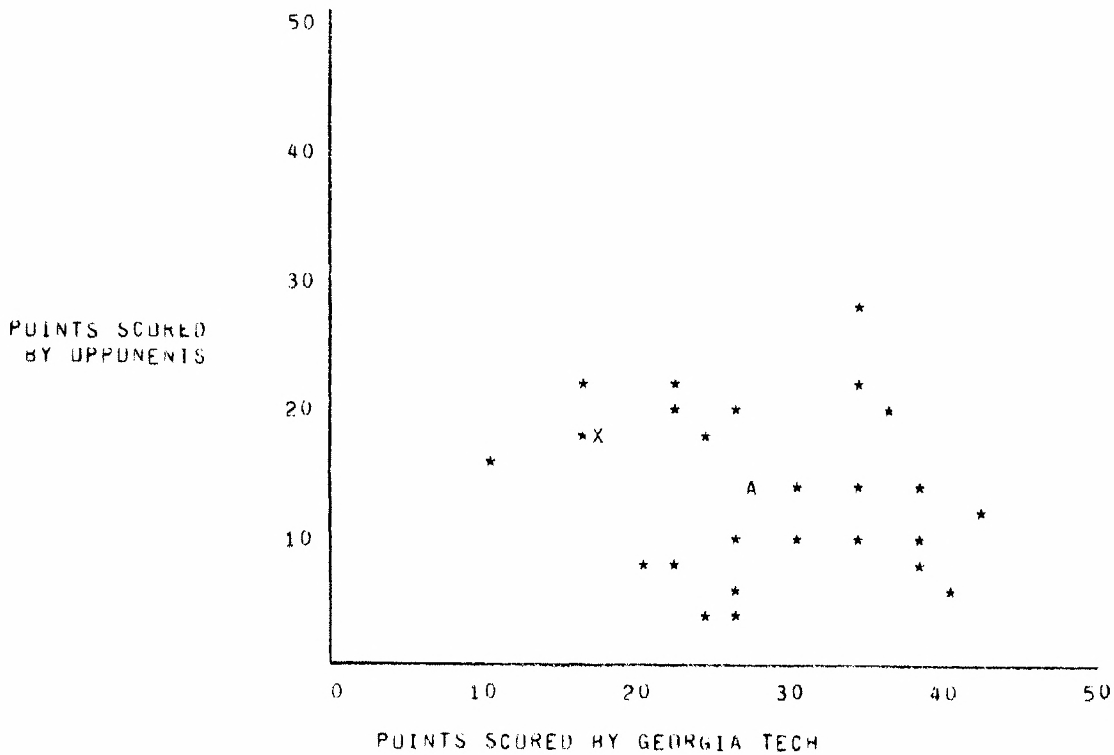


X ACTUAL SCORE
A SIMULATION AVG
* SIMULATED SCORE

	X	A	*																								
GEORGIA TECH	0	9	5	10	21	18	10	3	10	6	7	10	8	10	14	10	3	20	3	6	0	3	0	16	12	10	0
OPPONENTS	7	9	16	15	10	6	13	20	6	6	13	3	3	13	6	7	16	0	19	6	6	13	7	7	6	3	10

Figure 34. Output of the Simulation of the Georgia Tech-Georgia Game, 1964

FLORIDA 1960



X ACTUAL SCORE
A SIMULATION AVG
* SIMULATED SCORE

	X	A	*																								
GEORGIA TECH	17	28	34	37	26	26	16	33	23	29	22	37	26	33	40	22	34	21	25	41	9	19	16	23	30	37	35
OPPONENTS	18	13	10	13	6	3	18	21	17	14	22	7	19	27	6	7	13	20	10	12	16	7	22	3	10	9	19

Figure 35. Output of the Georgia Tech-Florida Game of 1960 with Georgia Tech's Decision Rule Modified so that Georgia Tech Does Not Kick on Third Down

BIBLIOGRAPHY

1. Adlinger, G. R., "Business Games--Play One!" *Harvard Business Review*, March-April, 1958.

2. Brosnan, J., "The New Books," *Harper's*, October, 1966, pp. 111-112.

3. Brownlee, K. A., *Statistical Theory and Methodology in Science and Engineering*, John Wiley and Sons, Inc., New York, 1965.

4. Cook, E., and Garner, W., *Percentage Baseball*, M. I. T. Press, Cambridge, Massachusetts, 1966.

5. Duesenberry, J. S., et al., *The Brookings-SSRC Quarterly Econometric Model of the United States*, Rand-McNally and North-Holland Press, Chicago, 1965.

6. Forrester, J., *Industrial Dynamics*, The M. I. T. Press and John Wiley and Sons, New York, 1961.

7. Hammersly, J. M., and Handscomb, D. C., *Monte Carlo Methods*, Spottiswoode, Ballantyne and Co., Ltd., 1964.

8. Hollingsdale, S. H. (Ed.), *Digital Simulation in Operational Research*, American Elsevier Publishing Company, Inc., New York, 1967.

9. Horowitz, I., "Probability Model for Baseball Management," *Journal of Industrial Engineering*, July-August, 1963, pp. 163-70.

10. Kibbee, J. M., et al., *Management Games*, Rheinhold Publishing Co., New York, 1961.

11. Meyer, H. A. (Ed.), *Symposium on Monte Carlo Methods*, John Wiley and Sons, Inc., New York, 1956.

12. Naylor, T., et al., *Computer Simulation Techniques*, John Wiley and Sons, Inc., New York, 1966.

13. Orcutt, G., et al., *Microanalysis of Socioeconomic Systems: A Simulation Study*, Harper and Brothers, New York, 1961.

14. Rapoport, A., *Strategy and Conscience*, Harper and Row, New York, 1964.

15. Ricciardi, F. M., et al., *Top Management Decision Making*, American Management Association, New York, 1957.

16. Siegel, S., *Nonparametric Statistics for the Behavioral Sciences*, McGraw-Hill, New York, 1956.

17. Tocher, K., *The Art of Simulation*, Van Nostrand, Princeton, N. J., 1963.

18. Wallace, W. N., "Pro Football Going Electronic," *The New York Times*, October 22, 1967, p. 2-5.

19. International Business Machines Corporation, *Bibliography on Simulation*, White Plains, N. Y., 1966.

20. National Collegiate Athletic Service, *Official Football Guide*, New York, 1960, 1961, 1962, 1963, 1964.